RIACS

Summation by Parts, Projections, and Stability

Pelle Olsson

Summation by Parts, Projections, and Stability

Pelle Olsson

Abstract

We have derived stability results for high-order finite difference approximations of mixed hyperbolic-parabolic initial-boundary value problems (IBVP). The results are obtained using summation by parts and a new way of representing general linear boundary conditions as an orthogonal projection. By rearranging the analytic equations slightly we can prove strict stability for hyperbolic-parabolic IBVP. Furthermore, we generalize our technique so as to yield strict stability on curvilinear non-smooth domains in two space dimensions. Finally, we show how to incorporate inhomogeneous boundary data while retaining strict stability. Using the same procedure one can prove strict stability in higher dimensions as well.

1 Introduction

When solving a partial differential equation numerically it is necessary to have some bound of the growth rate of the solution, since otherwise round-off errors could grow arbitrarily fast. This upper bound can be established by ensuring some kind of stability. We have elected to use the energy method, because it can be applied to the continuous as well as the discrete model. Furthermore, it can be applied to general domains, which is important when studying multidimensional problems.

Stability of the continuous problem is established by means of an integration-by-parts procedure introducing boundary terms, some of which must be eliminated to ensure stability. For the finite difference model integration by parts is replaced by summation by parts. This amounts to designing the discrete difference operator ensuring that, in addition to the accuracy requirements, certain conditions of antisymmetry are met. As a consequence, the common problem of finding proper "numerical" boundary conditions will be eliminated; they will be built in the discrete difference operator.

The analytic boundary conditions are yet to be incorporated. We propose a certain projection operator, which interacts with the difference operator so as to generate boundary terms that are completely analogous to those of the continuous problem. This can be done for any type of linear boundary conditions. Thus, an energy estimate is obtained for the discrete problem, provided there is one for the analytic model. This conclusion remains true for domains in several space dimensions, even if the boundary is non-smooth. Furthermore, using this projection operator allows us to derive stability results for a larger class of finite difference operators than those considered in [5]. Stability will be proved for high-order finite difference approximations of mixed hyperbolic-parabolic variable coefficient systems subject to general inhomogeneous boundary conditions.

1.1 An Introductory Example

To illustrate the underlying principles of the energy method we consider the convection-diffusion equation

$$u_t = u_{xx} + u_x, \quad x \in (0,1) \ t > 0$$
$$u(x,0) = f(x)$$
$$u(0,t) = 0$$
$$u_x(1,t) = g(t)$$

In the sequel we shall use the standard L^2-scalar product

$$(u,v) = \int_0^1 uv\,dx$$

with the corresponding norm defined as $||u||^2 = (u,u)$.

We can obtain an a priori estimate for this example using the following tools.

(i) Integration by parts:

$$\frac{d}{dt}||u||^2 = 2(u, u_{xx}) + 2(u, u_x) = -2||u_x||^2 + 2(u, u_x) + 2uu_x|_0^1$$

(ii) Boundary conditions:

$$\frac{d}{dt}||u||^2 = -2||u_x||^2 + 2(u, u_x) + 2u(1,t)g(1,t)$$

(iii) Cauchy-Schwarz inequality:

$$\frac{d}{dt}||u||^2 \leq -2||u_x||^2 + 2||u||||u_x|| + 2u(1,t)g(1,t)$$

(iv) Algebraic inequality:
$$2|xy| \leq \epsilon x^2 + \epsilon^{-1}y^2$$

implies ($\epsilon = 1$)

$$\frac{d}{dt}||u||^2 \leq -||u_x||^2 + ||u||^2 + u(1,t)^2 + g(1,t)^2$$

(v) Sobolev inequality:
$$|u|_\infty^2 \leq \epsilon||u_x||^2 + (\epsilon^{-1} + 1)||u||^2$$

is used to eliminate $u(1,t)$ ($\epsilon = 1$)

$$\frac{d}{dt}||u||^2 \leq 3||u||^2 + g(1,t)^2$$

which can be solved analytically to yield

$$||u(\cdot, t)||^2 \leq e^{3t}\left(||f||^2 + \int_0^t g(\tau)^2 d\tau\right)$$

If we are to obtain such an estimate for a system of equations we will also need

(vi) The adjoint of A:
$$(u, Av) = (A^T u, v)$$

Summing up, the energy method boils down to the six basic "tools" above. In the subsequent sections we shall see how these principles can be modified so as to give an energy estimate for the semi-discrete system.

2

2 General Principles for the Semi-discrete Case

In this section the basic principles of the energy method will be transferred to the semi-discrete case. Furthermore, a number of lemmas, which will be needed later, will be proved. Throughout this section grid vectors will be denoted by $v^T = (v_0^T \ldots v_\nu^T)$, $v_j \in \mathbb{R}^d$. Difference operators approximating $\partial/\partial x$ will be designated by

$$D = \frac{1}{h} \begin{pmatrix} d_{00}I & \ldots & d_{0\nu}I \\ \vdots & & \vdots \\ d_{\nu 0}I & \ldots & d_{\nu\nu}I \end{pmatrix} \quad I \in \mathbb{R}^{d\times d}$$

where D is written as a square matrix for convenience; in reality D will be a banded matrix, where the bandwidth is independent of the mesh size $h = L/\nu$.

2.1 Summation by Parts

In the semi-discrete case we employ summation by parts instead of integration by parts. The basic idea is to use difference operators satisfying

$$(u, Dv)_h = u_\nu^T v_\nu - u_0^T v_0 - (Du, v)_h \tag{1}$$

with respect to a *weighted* scalar product

$$(u, v)_h = h \sum_{i,j=0}^{\nu} \sigma_{ij} u_i^T v_j$$

It should be remarked that the usual Euclidean scalar product cannot be used. To prove the existence of summation by parts, it suffices to consider scalar products on the form

$$\Sigma = \begin{pmatrix} \Sigma^{(1)} & & \\ & I & \\ & & \Sigma^{(2)} \end{pmatrix}, \quad \Sigma^{(l)} \in \mathbb{R}^{(r_l+1)d \times (r_l+1)d}, \quad l = 1,2 \tag{2}$$

where the blocks of Σ are given by $\Sigma_{ij} = \sigma_{ij}I, I \in \mathbb{R}^{d\times d}$; r_l and the elements of $\Sigma^{(l)}$, $l = 1,2$ are independent of h. The following existence proof can be found in [5].

Proposition 2.1 *There exist scalar products (2) and difference operators D of accuracy $2p - 1$ at the boundaries and $2p$ in the interior, $p > 0$, such that the summation-by-parts property (1) holds.*

Confining ourselves to the case where $\Sigma^{(1)}$ and $\Sigma^{(2)}$ are diagonal we have the following existence theorem [4].

Proposition 2.2 *There exist diagonal scalar products (2) and difference operators D of accuracy p at the boundaries and $2p$ in the interior, $1 \leq p \leq 4$, such that the summation-by-parts property (1) holds.*

Remark: If one omits the requirement that the boundary stencils be at least accurate of order p for a given interior accuracy $2p$, it is possible to prove summation by parts for diagonal scalar products and difference operators D of arbitrary order of accuracy [7]. For a given boundary accuracy p, however, it may be necessary to resort to interior stencils of accuracy $q \gg 2p$, which may render these operators useless in practice.

The actual computation of the operators above is ill-conditioned, since it involves the solution of a rank-deficient problem. Using a symbolic language it is possible to solve for D exactly, the elements of which in general will depend on one or more parameters. Explicit examples can be found in [6]. For details on the algorithms we refer to [8]. The simplest example is furnished by

$$
D = \frac{1}{h}
\begin{pmatrix}
-1 & 1 & & & \\
-0.5 & 0 & 0.5 & & \\
& \ddots & \ddots & \ddots & \\
& & -0.5 & 0 & 0.5 \\
& & & -1 & 1
\end{pmatrix}
\tag{3}
$$

with the corresponding scalar product

$$
\Sigma = h
\begin{pmatrix}
0.5 & & & & \\
& 1 & & & \\
& & \ddots & & \\
& & & 1 & \\
& & & & 0.5
\end{pmatrix}
\tag{4}
$$

Summation by parts can be generalized to several space dimensions if we restrict ourselves to diagonal norms. To simplify the notation we consider only the two-dimensional case. A general proof is given in [6]. The grid function u_{ij} is partitioned as $u^T = (u_0^T \ldots u_{\nu_2}^T)$, $u_j^T = (u_{0j}^T \ldots u_{\nu_1 j}^T)$, $j = 0, \ldots, \nu_2$. Define the weighted scalar product as

$$
(u, v)_h = h \sum_{i=0}^{\nu} \sum_{j=0}^{\nu} \sigma_i \sigma_j u_{ij}^T v_{ij}
\tag{5}
$$

where we have assumed the same number of grid points in both dimensions for convenience only; $h = h_1 h_2$ is the cell area. Let D_1 and D_2 denote the difference operators approximating $\partial/\partial x_1$ and $\partial/\partial x_2$. Define

$$
(D_1 u)_{ij} = \frac{1}{h_1} \sum_{k=0}^{\nu} d_{ik} u_{kj} \qquad (D_2 u)_{ij} = \frac{1}{h_2} \sum_{k=0}^{\nu} d_{jk} u_{ik}
\tag{6}
$$

4

where it is assumed that the σ's and d's satisfy (1). Hence

$$(u, D_1 v)_h = h_2 \sum_{j=0}^{\nu} \sigma_j \left(\sum_{i=0}^{\nu} \sigma_i u_{ij}^T \sum_{k=0}^{\nu} d_{ik} v_{kj} \right)$$

and a similar expression holds for $(u, D_2 v)_h$. The parenthetical expression satisfies (1) for each j. We thus arrive at

Proposition 2.3 *Let the discrete difference operators D_1 and D_2 be defined by (6). Summation by parts then holds in both dimensions*

$$(u, D_1 v)_h = h_2 \sum_{j=0}^{\nu} \sigma_j u_{\nu j}^T v_{\nu j} - h_2 \sum_{j=0}^{\nu} \sigma_j u_{0j}^T v_{0j} - (D_1 u, v)_h$$

$$(u, D_2 v)_h = h_1 \sum_{i=0}^{\nu} \sigma_i u_{i\nu}^T v_{i\nu} - h_1 \sum_{i=0}^{\nu} \sigma_i u_{i0}^T v_{i0} - (D_2 u, v)_h$$

where $(\cdot, \cdot)_h$ is defined by (5).

Remark: This is the discrete counterpart of the two-dimensional divergence theorem. With a general domain Ω we assume that there is a smooth map $\xi = \xi(x)$ taking Ω onto the unit cube where proposition 2.3 can be applied. The assumption of such a map ξ is necessary in order for finite difference methods to apply to curvilinear domains. Consequently, integration by parts can always be replaced with summation by parts in the discrete case. It is presently unknown if it possible to obtain the summation-by-parts property in more than one dimension using non-diagonal norms.

2.2 Projections

Suppose that the model equation of section 1.1 were discretized as

$$\begin{aligned} v_t &= D^2 v + Dv \\ v(0) &= f \end{aligned} \tag{7}$$

where we have assumed homogeneous Neumann data for convenience; it will be shown later how to treat inhomogeneous boundary conditions. For every fixed h the problem above is a constant coefficient ODE system with a unique analytic solution. Consequently, there is little hope that the discretized boundary conditions $v_0(t) = (Dv)_\nu(t) = 0$ are fulfilled, since they have not been accounted for so far.

Denote by $V \subset \mathrm{R}^{\nu+1}$ the vector space where $v_0(t) = (Dv)_\nu(t) = 0$, and let P be a projection of v onto V. Multiplying (7) by P yields

$$(Pv)_t = P \left(D^2 v + Dv \right)$$

Any solution satisfying the boundary conditions must obey $v = Pv$, whence

$$v_t = P\left(D^2 v + Dv\right)$$

Conversely, we have

Proposition 2.4 *Let* $P \in R^{s \times s}$ *be a given projection independent of* t, *and suppose that* $v(t) \in R^s$ *is a solution of the non-linear ODE system*

$$\begin{aligned} v_t &= PR(t, v) + (I - P)g_t \\ v(0) &= f \end{aligned} \tag{8}$$

where f *satisfies* $f = Pf + (I - P)g(0)$. *Then*

$$v(t) = Pv(t) + (I - P)g(t), \quad t > 0$$

Proof:
Since P is independent of t, premultiplication of (8) gives ($P^2 = P$)

$$(Pv)_t = PR(t, v)$$

Using this equality in (8) implies

$$v_t = (Pv + (I - P)g)_t$$

Hence, by integration,

$$(I - P)(v(t) - g(t)) = (I - P)(f - g(0))$$

which proves the proposition. □

Remark: $g(t)$ represents the boundary data, and $(I - P)(v - g) = 0$ is the extension of $(I - P)v = 0$ to inhomogeneous boundary data. Proposition 2.4 thus tells us that any solution to (8) will satisfy the boundary conditions if the initial data do so.

In general P is not uniquely defined. Consider the vector space $V = \{v \in R^{\nu+1} | v_0 = 0, v_\nu = v_{\nu-1}\}$. Then

$$P = \begin{pmatrix} 0 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & 1 & 0 \end{pmatrix} \quad P = \begin{pmatrix} 0 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & & 1 \end{pmatrix}$$

both imply $Pv \in V$. To shed some light on how to choose P, we apply the energy method to (7)

$$\frac{d}{dt}\|v\|_h^2 = 2(v, P(D^2 v + Dv))_h$$

6

If P were self-adjoint w r t $(\cdot, \cdot)_h$, then

$$\frac{d}{dt}||v||_h^2 = 2(Pv, D^2v + Dv)_h = 2(v, D^2v + Dv)_h$$

where the last equality follows from proposition 2.4. The crucial condition to obtain this equality is expressed by

$$(u, Pv)_h = (Pu, v)_h \tag{9}$$

which states that P is an orthogonal projection (using the *weighted* scalar product $(\cdot, \cdot)_h$).

Suppose that $u(x, t) \in \mathbb{R}^d$, $x \in \mathbb{R}^n$ is a solution to

$$u_t = F(x, t, \partial)u \quad x \in \Omega$$
$$L(x, \partial)u = 0 \quad\quad x \in \Gamma$$

where ∂ denotes the n-dimensional gradient; Γ is the boundary of Ω. This system is discretized in space, possibly requiring a coordinate mapping onto the unit cube

$$v_t = PG(t, D)v$$

The projection P should be such that v fulfills

$$L^T v = 0$$

where L now represents a discretization of the analytic boundary conditions. Let $V = \{v \in \mathbb{R}^m | L^T v = 0\}$. According to the preceding discussion P is taken to be the orthogonal projection onto V (with respect to $(\cdot, \cdot)_h$). The boundary conditions can be written as

$$Q^T \Sigma v = 0$$

where $Q = \Sigma^{-1}L$. Hence, the boundary conditions are fulfilled for all vectors v that are orthogonal to the column space of Q, the orthogonal projection onto which reads $Q(Q^T\Sigma Q)^{-1}Q^T\Sigma$. In case $\Sigma = I$ this is the standard projection. The desired boundary projection is thus given by

$$P = I - Q(Q^T\Sigma Q)^{-1}Q^T\Sigma$$

or

$$P = I - \Sigma^{-1}L(L^T\Sigma^{-1}L)^{-1}L^T \tag{10}$$

Remark: In order for the projection to be well-defined the inverse of $L^T\Sigma^{-1}L$ must exist, which follows iff L has full column rank. The latter will follow from assumptions on the analytic boundary conditions (consistency arguments).

Proposition 2.5 *Suppose that L has full column rank, and let P be defined by (10). Then*

(i) $P^2 = P$

(ii) $\Sigma P = P^T \Sigma$

(iii) $v = Pv \iff L^T v = 0$

Proof:
All statements are immediate consequences of (10). $\qquad\qquad\qquad\qquad$ $\square$

Remark: The second statement of proposition 2.5 is equivalent to (9).

2.3 A Discrete Sobolev Inequality

As seen in section 1.1 it is necessary to have a Sobolev inequality. The following proposition shows that there is a discrete Sobolev inequality for the norms that we are interested in. We present it in a form suitable for proving *strict* stability.

Proposition 2.6 *Let* $|| \cdot ||_h$ *and* D *be defined by (2) and (1),respectively. Then*

$$|v|_\infty^2 \leq \epsilon ||Dv||_h^2 + \left(\epsilon^{-1} + 1 + \mathcal{O}(h)\right) ||v||_h^2$$

where $\epsilon > 0$.

Proof:
Choose k, l such that
$$|v_k|^2 = \min_j \left(|v_j|^2\right)$$
$$|v_l|^2 = \max_j \left(|v_j|^2\right) \equiv |v|_\infty^2$$

Eq. (2) implies that

$$||v||_h^2 \geq h \left(\lambda_1 |v^{(1)}|^2 + \lambda_2 |v^{(2)}|^2\right) + h \sum_{j=r_1+1}^{\nu-r_2} |v_j|^2$$

where $\lambda_{1,2} > 0$ are the smallest eigenvalues of $\Sigma^{(1,2)}$. Note that $\lambda_{1,2}$ are independent of h. Hence
$$||v||_h^2 \geq (1 - h\left(r_1(1 - \lambda_1) + r_2(1 - \lambda_2)\right)) |v_k|^2$$
where we have used $h\nu = L = 1$. If $c \equiv r_1(1 - \lambda_1) + r_2(1 - \lambda_2) \leq 0$ one immediately gets $|v_k|^2 \leq ||v||_h^2$. Otherwise we choose h such that $hc < 1$. Hence

$$|v_k|^2 \leq \frac{1}{1 - hc} ||v||_h^2 \leq (1 + Kh) ||v||_h^2 \qquad K = \frac{c}{1 - h_0 c} \qquad (11)$$

for $h \leq h_0$, where h_0 is a fixed number such that $h_0 c < 1$.

8

Next, we define a family of norms, which is obtained by shrinking the interior of (2); $\Sigma^{(1,2)}$ remain constant. Allowing a slight abuse of notation we write these norms as

$$(u, v)_{h,r,s} = \sum_{j=r}^{s} \sigma_{ij} u_i^T v_j$$

where $r \geq 0$ and $s \leq \nu$. Shrinking the interior of D accordingly one has

$$(v, Dv)_{h,k,l} = |v_l|^2 - |v_k|^2 - (Dv, v)_{h,k,l}$$

i. e.,

$$|v|_\infty^2 \leq |v_k|^2 + 2||Dv||_{h,k,l}||v||_{h,k,l}$$

Obviously $||v||_{h,k,l} \leq ||v||_{h,0,\nu} \equiv ||v||_h$, whence

$$|v|_\infty^2 \leq \epsilon||Dv||_h^2 + \left(\epsilon^{-1} + 1 + \mathcal{O}(h)\right)||v||_h^2$$

where (11) and the standard algebraic inequality have been used. $\square$

2.4   Adjoint Operators

As usual A^T denotes the transpose of A. We know from section 1.1 that $(u, Av) = (A^T u, v)$, i. e., the transpose of A is the adjoint operator. The question is whether A^T also is the adjoint operator with respect to $(\cdot, \cdot)_h$ as defined by (2). Let

$$A = \begin{pmatrix} A_0 & & \\ & \ddots & \\ & & A_\nu \end{pmatrix} \quad A_j = A(jh), \quad j = 0, \ldots, \nu, \quad h\nu = 1 \qquad (12)$$

denote the matrix representation of $A(x) \in \mathrm{R}^{d \times d}$, $x \in [0, 1]$. Smoothness will be assumed as needed.

Lemma 2.1 *Let Σ and A be defined by (2) and (12), respectively. Then*

$$|(u, Av)_h - (A^T u, v)_h| \leq \mathcal{O}(h)||u||_h||v||_h$$

Proof:
Denote the commutator of Σ and A by $[\Sigma, A]$. Then

$$(u, Av)_h = (A^T u, v)_h + h u^T [\Sigma, A] v \qquad (13)$$

where

$$[\Sigma, A] = \begin{pmatrix} [\Sigma^{(1)}, A^{(1)}] & & \\ & 0 & \\ & & [\Sigma^{(2)}, A^{(2)}] \end{pmatrix}$$

9

with

$$[\Sigma^{(1)}, A^{(1)}] = \begin{pmatrix} 0 & \sigma_{01}(A_1 - A_0) & \dots & \sigma_{0r}(A_r - A_0) \\ -\sigma_{01}(A_1 - A_0) & 0 & \dots & \sigma_{1r}(A_r - A_1) \\ \vdots & & & \vdots \\ -\sigma_{0r}(A_r - A_0) & -\sigma_{1r}(A_r - A_1) & \dots & 0 \end{pmatrix}$$

The other non-zero block has a similar structure. Assuming that $A(x)$ is differentiable we can apply the mean value theorem

$$[\Sigma^{(1)}, A^{(1)}] = h \begin{pmatrix} 0 & \sigma_{01} A'_{10} & \dots & \sigma_{0r} r A'_{r0} \\ -\sigma_{01} A'_{10} & 0 & \dots & \sigma_{1r}(r-1) A'_{r1} \\ \vdots & & & \vdots \\ -\sigma_{0r} r A'_{r0} & -\sigma_{1r}(r-1) A'_{r1} & \dots & 0 \end{pmatrix}$$

Hence,

$$|(u, Av)_h - (A^T u, v)_h| \le c |A'|_\infty h^2 \left(|u^{(1)}| |v^{(1)}| + |u^{(2)}| |v^{(2)}| \right) \le \mathcal{O}(h) \|u\|_h \|v\|_h$$

which proves the lemma. $\qquad\qquad\square$

Remark: According to lemma 2.1 the transpose of A is an approximate adjoint with respect to $(\cdot, \cdot)_h$, and the perturbation consists of lower order terms. The following assumption will be crucial when proving strict stability for hyperbolic systems.

Assumption 2.1 *Let A and Σ be given by (12) and (2). Then one of the conditions below is assumed to hold.*

   (i) Σ is diagonal $\text{diag}(\sigma_0 I \dots \sigma_\nu I)$.

   (ii) The blocks of A satisfy $A_0 = \dots = A_{r_1}$ and $A_{\nu - r_2 + 1} = \dots = A_\nu$.

Corollary 2.1 *If assumption 2.1 holds, then $(u, Av)_h = (A^T u, v)_h$.*

Proof:
In both cases we get $[\Sigma, A] = 0$. The result follows immediately from (13). $\qquad\square$

Remark: The latter criterion is satisfied if there is a $\delta > 0$ such that $A(x) = \text{const}$ for $0 \le x < \delta$ and $1 - \delta < x \le 1$, and if h is chosen such that $hr < \delta$, where $r = \max(r_1, r_2)$.

Corollary 2.2 *Let A be given by (12). If A is symmetric then then $(u, Au)_h = (Au, u)_h$.*

Proof:
A symmetric implies that $[\Sigma, A]$ is antisymmetric. Hence $u^T [\Sigma, A] u = 0$. $\qquad\square$

2.5 Some Operator Estimates

In this subsection we have gathered some operator estimates that will be needed in subsequent sections. The results are valid for norms defined by (2) unless otherwise stated. In particular, the estimates will be given in a form suitable for proving strict stability of the semi-discrete systems.

Lemma 2.2 *Let Σ and A be defined by (2) and (12). Then*

$$|(u, Av)_h| \leq |A|_\infty (1 + \mathcal{O}(h)) ||u||_h ||v||_h$$

where $|A|_\infty = \sup |A(x)|$.

Proof:
The definition of $(\cdot, \cdot)_h$ implies that $(u, Av)_h = h\tilde{u}^T \tilde{A} \tilde{v}$, where $\tilde{u} = \Sigma^{1/2} u$, $\tilde{v} = \Sigma^{1/2} v$, and $\tilde{A} = \Sigma^{1/2} A \Sigma^{-1/2}$. Taylor expansion yields $\tilde{A} = A + R$,

$$R = \begin{pmatrix} R^{(1)} & & \\ & 0 & \\ & & R^{(2)} \end{pmatrix}$$

with $R^{(l)} = \mathcal{O}(h)$, $l = 1, 2$. Thus

$$|(u, Av)_h| \leq |A|_\infty ||\tilde{u}|| ||\tilde{v}|| + \mathcal{O}(h) \left(||\tilde{u}^{(1)}|| ||\tilde{v}^{(1)}|| + ||\tilde{u}^{(2)}|| ||\tilde{v}^{(2)}|| \right)$$

i. e.,
$$|(u, Av)_h| \leq (|A|_\infty + \mathcal{O}(h)) ||\tilde{u}|| ||\tilde{v}||$$

where $|| \cdot ||$ denotes the standard Euclidean norm. Since $||\tilde{u}|| = ||u||_h$, $||\tilde{v}|| = ||v||_h$, the lemma follows. □

Corollary 2.3 *If, in addition to the hypotheses of lemma 2.2, assumption 2.1 is fulfilled, then*
$$|(u, Av)_h| \leq |A|_\infty ||u||_h ||v||_h$$

Proof:
The hypotheses imply that $\tilde{A} = A$, and the corollary follows. □

Remark: Lemma 2.2 states that the growth rate induced by low order terms is the same (modulus $\mathcal{O}(h)$-terms) in the continuous and the semi-discrete case.

It is well-known that $(u, [D, A]v)_h \leq |||[D, A]|||_h ||u||_h ||v||_h$, where $|||[D, A]|||_h$ can be bounded independent of h. This result can be sharpened under certain circumstances.

Lemma 2.3 *Let D be a difference approximation satisfying the summation-by-parts rule (1) with respect to a weighted norm (2), and define A by (12). Suppose that assumption 2.1 holds. If A is symmetric, then*

$$(u, [D, A]v)_h \leq \rho([D, A])||u||_h||v||_h$$

Proof:
According to the definition of the operator norm we have

$$||[D, A]||_h^2 = \max_{||v||_h=1} ||[D, A]v||_h^2 = \max_{||w||=1} hw^T C^T C w$$

where $w = \Sigma^{1/2}v$, $C = \Sigma^{1/2}[D, A]\Sigma^{-1/2}$. Because of the assumptions on A (or Σ) we have $C = \tilde{D}A - A\tilde{D}$, where $\tilde{D} = \Sigma^{1/2}D\Sigma^{-1/2}$. Summation by parts implies that

$$\Sigma D = D_s + D_a, \quad D_s = \frac{1}{2}\begin{pmatrix} -I & & \\ & 0 & \\ & & I \end{pmatrix} \quad I \in \mathbb{R}^{d \times d}$$

and D_a is an anti-symmetric matrix. Consequently,

$$C = [\Sigma^{-1/2}D_a\Sigma^{-1/2}, A]$$

where we have used $[\Sigma^{-1/2}D_s\Sigma^{-1/2}, A] = 0$. Since D_a is anti-symmetric and A symmetric we have $C^T = C$, i. e.,

$$||[D, A]||_h^2 = \max_{||w||=1} hw^T C^2 w = \rho(C)^2$$

Finally, $C = \Sigma^{1/2}[D, A]\Sigma^{-1/2}$ implies that

$$||[D, A]||_h = \rho([D, A])$$

which proves the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3 Homogeneous Boundary Conditions in One Dimension

We shall successively consider hyperbolic, parabolic and mixed hyperbolic-parabolic systems. Variable coefficient matrices will be allowed. To simplify the presentation we shall only deal with the lower boundary $x = 0$, which is justified if we take the solution to have compact support. In general, the upper boundary $x = 1$ is treated in a fashion similar to the procedure at the lower boundary.

3.1 Hyperbolic Systems

Consider the hyperbolic system

$$
\begin{aligned}
u_t &= \Lambda u_x + Bu + F \quad x \in (0,1) \\
u(x,0) &= f(x) \\
u_-(0,t) &= Lu_+(0,t) \quad L \in \mathrm{R}^{d_1 \times d_2}
\end{aligned}
\qquad
\Lambda(x,t) = \begin{pmatrix} \Lambda_-(x,t) & \\ & \Lambda_+(x,t) \end{pmatrix}
\qquad (14)
$$

where $u \in \mathrm{R}^d$, $d_1 + d_2 = d$; Λ_-, Λ_+ is the partitioning of Λ into negative and positive eigenvalues. It is assumed that the elements of the diagonal matrix Λ never change sign at the boundaries $x = 0$ and $x = 1$, and that there is a constant $\gamma > 0$ such that $\Lambda_-(j,t) \le -\gamma$ and $\Lambda_+(j,t) \ge \gamma$, $j = 0,1$ This implies that the rank of L is constant. Furthermore, L is assumed to be "small".

The discrete boundary conditions are written as $L^T v = 0$, where

$$
L^T = \begin{pmatrix} L_0^T & 0 & \dots & 0 \end{pmatrix} \in \mathrm{R}^{d_1 \times (\nu+1)d}
\qquad (15)
$$

Here $L_0^T = \begin{pmatrix} I & -L \end{pmatrix} \in \mathrm{R}^{d_1 \times d}$, the latter L being the analytic boundary operator. It follows immediately that $\mathrm{rank}(L) = \mathrm{rank}(I) = d_1$. The hypothesis of proposition 2.5 is thus satisfied, and we have the semi-discrete system

$$
\begin{aligned}
v_t &= P(\Lambda D v + B v + F) \\
v(0) &= f
\end{aligned}
\qquad
\Lambda = \begin{pmatrix} \Lambda(0,t) & & \\ & \ddots & \\ & & \Lambda(1,t) \end{pmatrix}
\qquad (16)
$$

Proposition 3.1 *Let $(\cdot,\cdot)_h$ be given by (2) and suppose that D satisfies the conclusion of proposition 2.1. If P is defined by (10) and (15), then the solution of (16) satisfies an energy estimate*

$$
\|v(t)\|_h^2 + \int_0^t \left(|v_0(\tau)|^2 + |v_\nu(\tau)|^2 \right) d\tau \le K e^{(\alpha' + \mathcal{O}(h))t} \left(\|f\|_h^2 + \int_0^t \|F(\tau)\|_h^2 d\tau \right)
$$

Proof:
The energy method yields (using propositions 2.5, 2.4)

$$
\frac{d}{dt}\|v\|_h^2 = 2(v,v_t)_h = 2(v,P(\Lambda D v + B v + F))_h = 2(v,\Lambda D v)_h + 2(v,Bv)_h + 2(v,F)_h
$$

Summation by parts implies ($v_\nu = 0$)

$$
(v,\Lambda D v)_h = -v_0^T \Lambda_0 v_0 - (Dv, \Lambda v)_h - (v, [D,\Lambda]v)_h
$$

Hence, by lemma 2.1

$$
(v,\Lambda D v)_h \le -\frac{1}{2} v_0^T \Lambda_0 v_0 + \frac{1}{2} \left(K_0 \|v\|_h \|hDv\|_h + \|[D,\Lambda]\|_h \|v\|_h^2 \right)
$$

13

where hD is a *bounded* operator, i. e.,

$$(v, \Lambda D v)_h \le -\frac{1}{2} v_0^T \Lambda_0 v_0 + \frac{1}{2} \left(K_1 + \|[D, \Lambda]\|_h \right) \|v\|_h^2$$

Now, according to propositions 2.4, 2.5 we have $L^T v = 0$, which is equivalent to $v_- = L v_+$ (the latter L denoting the analytic boundary operator). Thus

$$v_0^T \Lambda_0 v_0 = v_-^T \Lambda_- v_- + v_+^T \Lambda_+ v_+ = v_+^T \left(\Lambda_+ + L^T \Lambda_- L \right) v_+ \ge \frac{\gamma}{2} |v_0|^2$$

where the last inequality follows from the boundary conditions and the assumptions on L and Λ. Note that the analytic problem would result in exactly the same inequality. Hence

$$(v, \Lambda D v)_h \le -\frac{\gamma}{4} |v_0|^2 + \frac{1}{2} \left(K_1 + \|[D, \Lambda]\|_h \right) \|v\|_h^2$$

Lemma 2.1 shows that

$$(v, B v)_h \le \left(|B|_\infty + \mathcal{O}(h) \right) \|v\|_h^2$$

Consequently,

$$\frac{d}{dt} \|v\|_h^2 + |v_0|^2 \le \frac{1}{\min(1, \gamma/2)} \left((\||[D, \Lambda]|\|_h + 2|B|_\infty + 1 + K_1 + \mathcal{O}(h)) \|v\|_h^2 + \|F\|_h^2 \right)$$

Integration with respect to t proves the proposition with $K = \max(1, 2/\gamma)$. $\square$

Definition 3.1 *A semi-discrete approximation to the initial-boundary value problem $u_t = F(x, t, \partial)u$ is said to be strictly stable, if the semi-discrete solution satisfies an energy estimate that is exponentially bounded by $\exp(\alpha' t)$, $\alpha' = \alpha + O(h)$, where α is the expontial growth factor of the analytic estimate.*

Remark: If Λ (or $(\cdot, \cdot)_h$) satisfies the assumption 2.1, it follows that $K_1 = 0$. Also, by lemma 2.3, $\||[D, \Lambda]|\|_h = \rho([D, \Lambda])$. Eq. (16) would thus be strictly stable if $\rho([D, \Lambda]) \le |\Lambda'|_\infty$. In particular, (16) is strictly stable if $\Lambda(x) = $ const, since this implies $[D, \Lambda] = 0$. We also point out that the proportionality constant K is completely independent of the discretization. In case the estimate of the boundary integral is not needed one may take $K = 1$. For variable coefficient problems we have the following result:

Corollary 3.1 *Let D and $(\cdot, \cdot)_h$ be given by (3) and (4). Then (16) is strictly stable.*

Proof:
According to the preceding remark the corollary follows if we can show that $\rho([D, \Lambda]) \le |\Lambda'|_\infty$. But

$$[D, \Lambda] = \frac{1}{h} \begin{pmatrix} 0 & \Lambda_1 - \Lambda_0 & & & \\ 0.5(\Lambda_1 - \Lambda_0) & 0 & 0.5(\Lambda_2 - \Lambda_1) & & \\ & \ddots & \ddots & \ddots & \\ & & 0.5(\Lambda_{\nu-1} - \Lambda_{\nu-2}) & 0 & 0.5(\Lambda_\nu - \Lambda_{\nu-1}) \\ & & & \Lambda_\nu - \Lambda_{\nu-1} & 0 \end{pmatrix}$$

14

Assuming that $\Lambda(x)$ is C^1 the mean value theorem gives $\Lambda_i - \Lambda_j = \Lambda'(\xi_{ij})(i-j)h$ for some $\xi_{ij} \in (ih, jh)$. The corollary thus follows from the Gersgorin disc theorem. $\qquad\square$

3.2 Parabolic Systems

We consider the parabolic system

$$
\begin{aligned}
u_t &= Au_{xx} + Bu_x + Cu + F \quad x \in (0,1)\\
u(x,0) &= f(x)\\
L_0 u(0,t) &+ L_1 u_x(0,t) = 0
\end{aligned}
\qquad
L_0 = \begin{pmatrix} L_0^I \\ L_0^{II} \end{pmatrix} \quad L_1 = \begin{pmatrix} L_1^I \\ 0 \end{pmatrix}
\qquad (17)
$$

where $L_{0,1}^I \in R^{d_1 \times d}$, $L_0^{II} \in R^{d_2 \times d}$, $d_1 + d_2 = d$; $\mathrm{rank}(L_1^I) = d_1$, $\mathrm{rank}(L_0^{II}) = d_2$; A, B, C, and F depend smoothly on x and t. It is assumed that the system is strongly parabolic, i. e., $A(x,t) + A(x,t)^T \geq 2\delta I$.

The following lemma, a proof of which can be found in [3], will be crucial when proving an energy estimate for the solution of (17) and its semi-discrete counterpart.

Lemma 3.1 *Let $A \in R^{d \times d}$ be arbitrary and let $L_0, L_1 \in R^{d \times d}$ be of the form (17). The following conditions are equivalent:*

(i) There exists a constant $c > 0$ such that

$$|u^T A u_x| \leq c|u|^2$$

for all u, $u_x \in R^d$ that satisfy

$$L_0 u + L_1 u_x = 0$$

(ii) If a, $b \in R^d$ are vectors such that

$$L_1^I b = 0, \quad L_0^{II} a = 0$$

then

$$a^T A b = 0$$

Assumption 3.1 *Given the boundary matrices L_0, L_1, the matrix A is supposed to be such that the second condition of lemma 3.1 holds.*

Remark: Except for Dirichlet and Neumann conditions, assumption 3.1 imposes severe restrictions on A. Lemma 3.1 states that the assumption above is necessary in order to obtain an energy estimate. The computations that follow will show how the second condition, which holds by assumption, implies the first.

Before deriving the energy estimates, one more lemma is needed [3].

Lemma 3.2 *Suppose that assumption 3.1 holds and that $A(x,t) + A(x,t)^T \geq 2\delta I$. Then the $d \times d$ matrix*

$$\begin{pmatrix} L_1^I \\ L_0^{II} \end{pmatrix}$$

is non-singular.

As usual the boundary conditions are written as $L^T v = 0$, where

$$L^T = \begin{pmatrix} L_0 + \dfrac{d_{00}}{h}L_1 & \dfrac{d_{01}}{h}L_1 & \ldots & \dfrac{d_{0r}}{h}L_1 & 0 & \ldots 0 \end{pmatrix} \in \mathrm{R}^{d \times (\nu+1)d} \tag{18}$$

where d_{0j}/h are the non-zero elements of the first row of D, which is a difference operator satisfying the conclusion of proposition 2.1 or 2.2. We have

$$L_0 + \frac{d_{00}}{h}L_1 = \begin{pmatrix} (d_{00}/h)I^I & \\ & I^{II} \end{pmatrix} \left[\begin{pmatrix} L_1^I \\ L_0^{II} \end{pmatrix} + \frac{h}{d_{00}} \begin{pmatrix} L_0^I \\ 0 \end{pmatrix} \right]$$

Thus, lemma 3.2 implies that $L_0 + (d_{00}/h)L_1$ is non-singular for $h > 0$ sufficiently small. From (18) it follows immediately that $\mathrm{rank}(L) = d$. According to proposition 2.5 the corresponding projection operator is well-defined, and we obtain

$$\begin{aligned} v_t &= P(AD^2 v + BDv + Cv + F) \\ v(0) &= f \end{aligned} \qquad A = \begin{pmatrix} A(0,t) & & \\ & \ddots & \\ & & A(1,t) \end{pmatrix} \tag{19}$$

with similar expressions for B, C, F.

Proposition 3.2 *Let $(\cdot, \cdot)_h$ be given by (2) and suppose that D satisfies the conclusion of proposition 2.1. If P is defined by (10) and (18), then the solution of (19) satisfies an energy estimate*

$$||v(t)||_h^2 + \int_0^t \left(|v_0(\tau)|^2 + |v_\nu(\tau)|^2 \right) d\tau \leq e^{(\alpha' + \mathcal{O}(h))t} \left(||f||_h^2 + \int_0^t ||F(\tau)||_h^2 d\tau \right)$$

Proof:
By propositions 2.5, 2.4, 2.1 we have

$$(v, PAD^2 v)_h = (v, AD^2 v)_h = -v_0^T A(Dv)_0 - (Dv, ADv)_h - (v, [D, A]Dv)_h$$

where we have assumed homogeneous Dirichlet conditions at the upper boundary for convenience. Due to proposition 2.5 it follows that

$$L_0 v_0 + L_1 \frac{1}{h} \sum_{j=0}^{r} d_{0j} v_j = L_0 v_0 + L_1 (Dv)_0 = 0 \tag{20}$$

16

Partition $v_j = v_j' + v_j''$, $v_j' \in \ker L_1^I$, $v_j'' \in (\ker L_1^I)^\perp$. Eq. (20) implies $L_0^{II} v_0 = 0$ and by construction $L_1^I (Dv')_0 = 0$. Hence, according to assumption 3.1

$$-v_0^T A(Dv)_0 = -v_0^T A(Dv'')_0$$

Eq. (20) can be rewritten as

$$\begin{pmatrix} L_1^I \\ 0 \end{pmatrix} (Dv'')_0 = -L_0 v_0$$

Since $(Dv'')_0 \in (\ker L_1^I)^\perp$ we get

$$\tilde{L}_1 (Dv'')_0 = -L_0 v_0, \quad \tilde{L}_1 = \begin{pmatrix} L_1^I \\ s_1^T \\ \vdots \\ s_{d_2}^T \end{pmatrix}$$

where $\{s_j\}$ is a basis in $\ker L_1^I$. Thus, $\tilde{L}_1$ is non-singular, and one obtains

$$-v_0^T A(Dv)_0 = v_0^T A \tilde{L}_1^{-1} L_0 v_0 \le \gamma |v_0|^2, \quad \gamma = |A \tilde{L}_1^{-1} L_0|_\infty$$

This is exactly the same expression as one would get in the analytic case. Thus

$$(v, PAD^2 v)_h \le \gamma |v_0|^2 - \delta ||Dv||_h^2 + ||[D,A]||_h ||v||_h ||Dv||_h \tag{21}$$

Furthermore,

$$(v, PBDv)_h \le (|B|_\infty + \mathcal{O}(h)) \, ||v||_h ||Dv||_h, \quad (v, PCv)_h \le (|C|_\infty + \mathcal{O}(h)) \, ||v||_h^2 \tag{22}$$

Finally, proposition 2.6 and the algebraic inequality yield

$$\frac{d}{dt} ||v||_h^2 + |v_0|^2 \le (\alpha' + \mathcal{O}(h)) \, ||v||_h^2 + ||F||_h^2$$

Integration with respect to time proves the proposition. $\qquad\square$

Remark: All coefficients, except $||[D,A]||_h$, appearing in (21) and (22) are identical (modulus $\mathcal{O}(h)$-terms) to those of the analytic estimate. Since the discrete Sobolev inequality 2.6 introduces the same growth rate as the analytic Sobolev inequality, it follows that (19) is strictly stable if we have the estimate $||[D,A]||_h \le |A'|_\infty$, which is true if $A(x) = \text{const}$. For variable coefficients one can prove

Corollary 3.2 *Let D and $(\cdot, \cdot)_h$ be given by (3) and (4). Then (19) is strictly stable if A is symmetric.*

Proof:
Same as for corollary 3.1. $\qquad\square$

3.3 Hyperbolic-Parabolic Systems

Consider the mixed hyperbolic-parabolic system

$$
\begin{aligned}
u_t &= Au_{xx} + B_{11}u_x + B_{12}v_x + C_{11}u + C_{12}v + F \quad && x \in (0,1) \\
v_t &= \Lambda v_x + B_{21}u_x + C_{21}u + C_{22}v + G \quad && u \in \mathrm{R}^{d_1}, v \in \mathrm{R}^{d_2}
\end{aligned}
$$

$$
\begin{aligned}
L_1 u_x(0,t) + L_0 u(0,t) + M_0 v(0,t) &= 0 \\
v_-(0,t) &= S_0 v_+(0,t) + R_0 u(0,t)
\end{aligned}
\qquad v_- \in \mathrm{R}^{d_2'}, v_+ \in \mathrm{R}^{d_2''} \quad M_0 = \begin{pmatrix} M_0^I \\ 0 \end{pmatrix}
$$

$$
\begin{aligned}
u(x,0) &= f(x) \\
v(x,0) &= \phi(x)
\end{aligned}
$$

$$(23)$$

As usual we assume $u = v \equiv 0$ in a neighborhood of $x = 1$ for convenience; L_0, L_1 are as in section 3.2, and S_0 satisfies the hypotheses of the boundary operator in section 3.1. The coefficient matrices and the forcing functions of the differential equations may depend on x and t.

The discretized boundary conditions are written as $L^T w = 0$, where $L^T \in \mathrm{R}^{d' \times (\nu+1)d}$, $d = d_1 + d_2$, $d' = d_1 + d_2'$ is given by

$$
\left(
\begin{pmatrix}
L_0 + \dfrac{d_{00}}{h}L_1 & M_0 \\
-R_0 & \begin{pmatrix} I & -S_0 \end{pmatrix}
\end{pmatrix}
\begin{pmatrix}
\dfrac{d_{01}}{h}L_1 & 0 \\
0 & 0
\end{pmatrix}
\cdots
\begin{pmatrix}
\dfrac{d_{0r}}{h}L_1 & 0 \\
0 & 0
\end{pmatrix}
\quad 0 \ \cdots \ 0
\right)
$$

$$(24)$$

We want to show that L^T has full rank. The first block of L^T can be rewritten as

$$
\begin{pmatrix} D(h) & 0 \\ 0 & I \end{pmatrix}
\left[
\begin{pmatrix} \tilde{L} & 0 & 0 \\ -R_0 & I & 0 \end{pmatrix}
+
\begin{pmatrix} (h/d_{00})\hat{L} & (h/d_{00})M_{00} & (h/d_{00})M_{01} \\ 0 & 0 & -S_0 \end{pmatrix}
\right]
$$

where

$$
D(h) = \begin{pmatrix} (d_{00}/h)I^I & \\ & I^{II} \end{pmatrix}
\qquad
\tilde{L} = \begin{pmatrix} L_1^I \\ L_0^{II} \end{pmatrix}
\qquad
\hat{L} = \begin{pmatrix} L_0^I \\ 0 \end{pmatrix}
$$

Since $\tilde{L}$ is invertible, it follows that

$$
\begin{pmatrix} \tilde{L} & 0 \\ -R_0 & I \end{pmatrix}
+
\begin{pmatrix} (h/d_{00})\hat{L} & (h/d_{00})M_{00} \\ 0 & 0 \end{pmatrix}
$$

is invertible, i. e., has full rank for $h > 0$ sufficiently small. The expression enclosed by the square brackets thus has linearly independent rows, which in turn implies that the first block of L^T has full rank. Hence, L has full rank, and the corresponding projection is well-defined.

The semi-discrete system is formulated as

$$
\begin{aligned}
w_t &= P\left(\tilde{A}\tilde{D}^2 w + \tilde{\Lambda}\tilde{D}w + \tilde{C}w + \tilde{F} \right) \\
w(0) &= \psi
\end{aligned}
\qquad
w_j = \begin{pmatrix} u_j \\ v_j \end{pmatrix}, \quad j = 0, \ldots, \nu
\qquad (25)
$$

18

where

$$\tilde{A} = \begin{pmatrix} \begin{pmatrix} A(0,t) & 0 \\ 0 & 0 \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} A(1,t) & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix}$$

$$\tilde{\Lambda} = \begin{pmatrix} \begin{pmatrix} B_{11}(0,t) & B_{12}(0,t) \\ B_{21}(0,t) & \Lambda(0,t) \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} B_{11}(1,t) & B_{12}(1,t) \\ B_{21}(1,t) & \Lambda(1,t) \end{pmatrix} \end{pmatrix}$$

$$\tilde{C} = \begin{pmatrix} \begin{pmatrix} C_{11}(0,t) & C_{12}(0,t) \\ C_{21}(0,t) & C_{22}(0,t) \end{pmatrix} & & & \\ & \ddots & & \\ & & \begin{pmatrix} C_{11}(1,t) & C_{12}(1,t) \\ C_{21}(1,t) & C_{22}(1,t) \end{pmatrix} \end{pmatrix}$$

The forcing function $\tilde{F}$ and the initial data ψ are defined analogously.

Proposition 3.3 *Let $(\cdot,\cdot)_h$ be given by (2) and suppose that $\tilde{D}$ satisfies the conclusion of proposition 2.1. If P is defined by (10) and (24), then the solution of (25) satisfies an energy estimate*

$$\|u(t)\|_h^2 + \|v(t)\|_h^2 + \sum_{j=0,\nu} \int_0^t \left(|u_j(\tau)|^2 + |v_j(\tau)|^2\right) d\tau$$

$$\leq K e^{(\alpha' + \mathcal{O}(h))t} \left(\|f\|_h^2 + \|\phi\|_h^2 + \int_0^t \left(\|F(\tau)\|_h^2 + \|G(\tau)\|_h^2\right) d\tau\right)$$

Proof: The energy method applied to (25) yields

$$\frac{d}{dt}\|w\|_h^2 = 2(w, P(\tilde{A}\tilde{D}^2 w + \tilde{\Lambda}\tilde{D}w + \tilde{C}w + \tilde{F}))_h = 2(w, (\tilde{A}\tilde{D}^2 w + \tilde{\Lambda}\tilde{D}w + \tilde{C}w + \tilde{F}))_h$$

Now

$$\begin{aligned} (w, \tilde{A}\tilde{D}^2 w)_h &= h \sum_{i,j=0}^{\nu} \sigma_{ij}(u_i^T v_i^T) \begin{pmatrix} A_j & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{h^2} \sum_{k,l=0}^{\nu} d_{jk} d_{kl} w_l \\ &= h \sum_{i,j=0}^{\nu} \sigma_{ij} u_i^T A_j \frac{1}{h^2} \sum_{k,l=0}^{\nu} d_{jk} d_{kl} u_l = (u, AD^2 u)_h \end{aligned}$$

where D is the difference operator of (19). The remaining terms are handled in a similar manner. One has

(i) $(w, \tilde{A}\tilde{D}^2 w)_h = (u, AD^2 u)_h$

(ii) $(w, \tilde{\Lambda}\tilde{D}w)_h = (u, B_{11}Du)_h + (u, B_{12}Dv)_h + (v, B_{21}Du)_h + (v, \Lambda Dv)_h$

(iii) $(w, \tilde{C}w)_h = (u, C_{11}u)_h + (u, C_{12}v)_h + (v, C_{21}u)_h + (v, C_{22}v)_h$

(iv) $(w, \tilde{F})_h = (u, F)_h + (v, G)_h$

For convenience we use the same symbol D to denote the difference operators acting on u and v. As far as the energy estimate is concerned, the hyperbolic-parabolic system has now been reduced to the previously treated hyperbolic and parabolic systems.

Items (iii) and (iv) consist only of lower order terms, and can be estimated using lemma 2.2. Thus, the coefficients of the estimates are identical to the corresponding analytic estimate (modulo $\mathcal{O}(h)$-terms). In item (ii) the potentially "dangerous" terms are those containing Dv. Using exactly the same technique as in the proof of proposition 3.1 we get

$$(v, \Lambda Dv)_h \leq -\frac{\gamma}{4}|v_0|^2 + |S_0^T \Lambda_- R_0|_\infty |v_0||u_0| + \frac{1}{2}|R_0^T \Lambda_- R_0|_\infty |u_0|^2$$
$$+ \frac{1}{2}\left(K_1 + ||[D,\Lambda]||_h\right)||v||_h^2$$

i. e., by means of the algebraic inequality

$$(v, \Lambda Dv)_h \leq -\frac{\gamma}{6}|v_0|^2 + \gamma_1|u_0|^2 + \frac{1}{2}\left(K_1 + ||[D,\Lambda]||_h\right)||v||_h^2$$

Furthermore,

$$(u, B_{12}Dv)_h \leq \frac{1}{2}|B_{12}|_\infty \left(\epsilon_1|v_0|^2 + \epsilon_1^{-1}|u_0|^2\right) + ||[D,B_{12}]||_h||u||_h||v||_h - (Du, B_{12}v)_h$$

Finally, in item (i) the term $(u, AD^2 u)_h$ is treated as in the proof of proposition 3.2, the only difference being that

$$-u_0^T A(Du)_0 = u_0^T A\tilde{L}_1^{-1} L_0 u_0 + u_0^T A\tilde{L}_1^{-1} M_0 v_0 \leq \gamma_2 \left[\epsilon_2|v_0|^2 + \left(\epsilon_2^{-1} + 1\right)|u_0|^2\right]$$

We point out that the coefficients of the boundary terms in the inequalities above are identical to those of the analytic estimate. Choosing ϵ_1 and ϵ_2 sufficiently small we thus arrive at

$$\frac{d}{dt}||w||_h^2 + \frac{\gamma}{4}\left(|u_0|^2 + |v_0|^2\right) \leq (\alpha' + \mathcal{O}(h))||w||_h^2 + ||F||_h^2 + ||G||_h^2$$

where we have used $||w||_h^2 = ||u||_h^2 + ||v||_h^2$; in the right member we have used proposition 2.6 and the algebraic inequality to eliminate $|u_0|^2$ and $||Du||_h$. Integration proves the proposition with $K = \max(1, 4/\gamma)$. $\qquad \Box$

Remark: In case no estimate of $|v_0|^2$ is needed one may take $K = 1$. Also, only the coefficients $|||[D, A]|||_h$, $|||[D, \Lambda]|||_h$, $|||[D, B_{12}]|||_h$ and K_1 will be larger than their analytic counterparts. If either of the conditions of assumption 2.1 is met, then $K_1 = 0$ and the operator norms can be replaced by the corresponding spectral radii (cf. lemma 2.3). In particular, if A, Λ, B_{12} are constant, then (25) is strictly stable. As before, for variable coefficients we have

Corollary 3.3 *Let $\tilde{D}$ and $(\cdot, \cdot)_h$ be given by (3) and (4). Then (25) is strictly stable if A and B_{12} are symmetric.*

Proof:
Same as for corollary 3.1. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

3.4 Strict stability

So far we have obtained strict stability under special circumstances, such as constant coefficient problems or second order methods. The crux of the matter lies in estimating the commutator $[D, A]$. Only in the previous cases were we able to prove that $|||[D, A]|||_h \leq |A'|_\infty$. In fact, numerical experiments show that $|||[D, A]|||_h \geq \rho([D, A]) = K|A'|_\infty$, $K > 1$, for high-order methods. Typical values for D's corresponding to diagonal norms are $K = 1.67$, $K = 2.55$, and $K = 35.8$, where the operator accuracy increases from three to five. One would still obtain $K > 1$ even if one considered only the interior operator. This indicates that the commutator should be avoided, which can be achieved if the analytic problem is reformulated.

The hyperbolic system (14) can be rewritten in skew-symmetric form as

$$
\begin{aligned}
u_t &= \frac{1}{2}(\Lambda u)_x + \frac{1}{2}\Lambda u_x + \left(B - \frac{1}{2}\Lambda'\right)u + F \quad x \in (0, 1)\\
u(x, 0) &= f(x)\\
u_-(0, t) &= Lu_+(0, t) \qquad\qquad\qquad\qquad L \in \mathbf{R}^{d_1 \times d_2}
\end{aligned}
$$

The corresponding semi-discrete system becomes

$$
\begin{aligned}
v_t &= P\left(\frac{1}{2}D\Lambda v + \frac{1}{2}\Lambda Dv + \left(B - \frac{1}{2}\Lambda'\right)v + F\right)\\
v(0) &= f
\end{aligned}
\tag{26}
$$

Proposition 3.4 *Let $(\cdot, \cdot)_h$ be given by (2) and suppose that D satisfies the conclusion of proposition 2.1. Define P by (10) and (15). If either Λ or Σ fulfills assumption 2.1, then (26) is strictly stable.*

Proof:
The energy method implies

$$\frac{d}{dt}||v||_h^2 = -v_0^T \Lambda_0 v_0 - (Dv, \Lambda v)_h + (v, \Lambda Dv)_h - (v, \Lambda' v)_h$$
$$+2(v, Bv)_h + 2(v, F)_h$$

The boundary terms are treated exactly as in the proof of proposition 3.1. Because of corollary 2.1 we have $(Dv, \Lambda v)_h = (v, \Lambda Dv)_h$. Thus, by lemma 2.2,

$$\frac{d}{dt}||v||_h^2 + \frac{\gamma}{2}|v_0|^2 \le (|\Lambda'|_\infty + 2|B|_\infty + 1 + \mathcal{O}(h))||v||_h^2 + ||F||_h^2$$

which is identical (neglecting $\mathcal{O}(h)$-terms) to the analytic estimate. $\square$

Remark: If Σ is diagonal, then the $\mathcal{O}(h)$-terms vanish identically (corollary 2.3).

The parabolic system (17) is altered in a slightly different manner. The modified system reads
$$u_t = (Au_x)_x + (B - A')u_x + Cu + F \quad x \in (0,1)$$
$$u(x,0) = f(x)$$
$$L_0 u(0,t) + L_1 u_x(0,t) = 0$$

which is discretized as

$$v_t = P\left(DADv + (B - A')Dv + Cv + F\right)$$
$$v(0) = f \tag{27}$$

Proposition 3.5 *Let $(\cdot, \cdot)_h$ be given by (2) and suppose that D satisfies the conclusion of proposition 2.1. If P is defined by (10) and (18), then (27) is strictly stable.*

Proof:
Left to the reader. $\square$

Finally, the mixed hyperbolic-parabolic system is reformulated as

$$u_t = (Au_x)_x + (B_{11} - A')u_x + (B_{12}v)_x + C_{11}u + (C_{12} - B_{12}')v + F \quad x \in (0,1)$$
$$v_t = \frac{1}{2}(\Lambda v)_x + \frac{1}{2}\Lambda v_x + B_{21}u_x + C_{21}u + (C_{22} - \frac{1}{2}\Lambda')v + G$$

where the initial data and the boundary conditions are identical to those of (23). In semi-discrete form we have

$$w_t = P\left(\tilde{D}\tilde{A}\tilde{D}w + \tilde{D}\tilde{\Lambda}w + \tilde{B}\tilde{D}w + (\tilde{C} - \tilde{\Lambda}')w + \tilde{F}\right)$$
$$w(0) = \psi \tag{28}$$

where

$$\tilde{\Lambda} = \left(\begin{array}{ccc} \begin{pmatrix} 0 & B_{12}(0,t) \\ 0 & \Lambda(0,t)/2 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & B_{12}(1,t) \\ 0 & \Lambda(1,t)/2 \end{pmatrix} \end{array} \right)$$

$$\tilde{B} = \left(\begin{array}{ccc} \begin{pmatrix} B_{11}(0,t) - A'(0,t) & 0 \\ B_{21}(0,t) & \Lambda(0,t)/2 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} B_{11}(1,t) - A'(1,t) & 0 \\ B_{21}(1,t) & \Lambda(1,t)/2 \end{pmatrix} \end{array} \right)$$

Proposition 3.6 *Let $(\cdot,\cdot)_h$ be given by (2) and suppose that $\tilde{D}$ satisfies the conclusion of proposition 2.1. Define P by (10) and (24). If either Λ or Σ fulfills assumption 2.1, then (28) is strictly stable.*

Proof:
Left to the reader. □

# 4 Homogeneous Boundary Conditions in Two Dimensions

The results of section 3 will now be generalized to two space dimensions. If the boundary is smooth, the original problem can be decomposed into two problems via a partition of unity, one of which is a Cauchy problem. The second problem is an initial-boundary value problem that is periodic in one space dimension, see figure below.
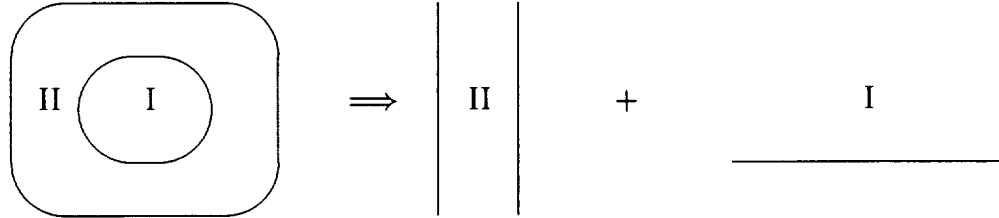
Consequently, summation by parts is needed only in one dimension, and the generalization of propositions 3.1, 3.2, 3.3 to two dimensions follows immediately. For details on the

decomposition we refer to [3]. The situation is different if the boundary is non-smooth, which is the case in the presence of corners. As mentioned at the end of section 2.1, it is not known how to extend norms of type (2) so as to obtain summation by parts in several space dimensions. We thus limit ourselves to diagonal norms, in which case we have proposition 2.3.

All boundary conditions considered so far are *local*. In case of characteristic and Dirichlet conditions no new difficulties are presented in two dimensions, because each boundary point can be treated individually. Boundary conditions involving derivatives increase the complexity significantly. Therefore, we shall only allow normal derivatives in the boundary operator. This is no serious restriction from the application point of view. Thus, away from the corners these boundary conditions are locally one-dimensional. For each such boundary point we obtain a projection operator of the previous section. In particular, these operators commute since they affect disjoint sets of grid points. At corners the situation is more complicated, because there are two different normal derivatives, which implies that the corresponding projection no longer is locally one-dimensional.

P_1, P_2, P_c commute
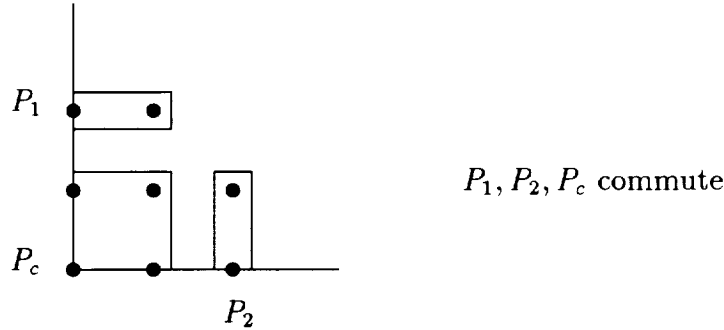
Throughout this section we shall focus our interest on the origin, and assume that the solutions are supported only in a neighborhood of $(0,0)$. The remaining boundary conditions will be accounted for by applying the projection operators corresponding to the boundary point in question. Since these operators commute, the resulting product is the uniquely defined boundary projection. The domain of definition is taken to be $\Omega = (0,1) \times (0,1)$ with boundary Γ. It will be shown later how to extend the results to curvilinear domains. In order to simplify the presentation all lower order terms will be omitted.

4.1 Symmetric Hyperbolic Systems

Consider

$$
\begin{aligned}
u_t &= \sum_{i=1}^{2} A_i u_{x_i} + F, && x \in \Omega = (0,1) \times (0,1), \quad u \in \mathbf{R}^d \\
u(x,0) &= f(x) && x = (x_1, x_2) \\
\varphi_I(x,t) &= S(x)\varphi_{II}(x,t), && x \in \Gamma
\end{aligned}
\tag{29}
$$

24

where φ_I, φ_{II} denote the locally ingoing and outgoing characteristic variables; $A_i = A_i(x,t)$, $i = 1, 2$ are symmetric and $S(x)$ is assumed to be "small". It should be noted that $\varphi_I \in \mathbb{R}^{d_1(x)}$, $\varphi_{II} \in \mathbb{R}^{d_2(x)}$, where $d_1(x) + d_2(x) = d$, $x \in \Gamma$. The matrix

$$A(x) \equiv \sum_{i=1}^{2} n_i(x) A_i(x) \qquad (30)$$

can be diagonalized for every $x \in \Gamma$; $n(x) = (n_1(x), n_2(x))$ is the outward unit normal of Γ. Hence

$$\Lambda(x) = Q^T(x) A(x) Q(x), \quad x \in \Gamma \qquad (31)$$

The characteristic variables are only needed at the boundary, and they are defined as $\varphi(x,t) = Q^T(x) u(x,t)$. It will be assumed that $\Lambda(x)$ is uniformly non-singular for $x \in \Gamma$, i. e., the eigenvalues are bounded away from zero. However, the number of positive and negative eigenvalues may differ from one boundary point to another. The analytic boundary conditions can thus be expressed as

$$L(x) u(x,t) = 0 \quad L(x) = \left(\begin{array}{cc} Q_I^T(x) & -S(x) Q_{II}^T(x) \end{array} \right) \qquad (32)$$

Clearly, $L(x)$ has full rank for every $x \in \Gamma$. Strictly speaking, $L(0,0)$ is not defined so far, because the normal $n(0,0)$ is not well-defined. It will soon be shown how to define $L(0,0)$, and we can formally consider $L(x)$ as being defined for every $x \in \Gamma$.

Let v_{ij}, $i = 0, \ldots, \nu_1$, $j = 0, \ldots, \nu_2$ be a grid function. Define $v^T = (v_0^T \ldots v_{\nu_2}^T)$, $v_j^T = (v_{0j}^T \ldots v_{\nu_1 j}^T)$. The discretized boundary conditions are written as

$$L_{ij}^T v_j = 0, \quad i = 0, \nu_1, \quad j = 1, \ldots, \nu_2 - 1 \text{ and } j = 0, \nu_2, \quad i = 0, \ldots, \nu_1 \qquad (33)$$

where

$$L_{ij}^T = \left(\begin{array}{ccccccc} 0 & \ldots & 0 & L(ih_1, jh_2) & 0 & \ldots & 0 \end{array} \right) \in \mathbb{R}^{d_1(i,j) \times (\nu_1+1)d}$$

with the non-zero element being the i th entry. At the origin we define

$$L(0,0) = \left(\begin{array}{cc} Q_I^T(0,0) & -S(0,0) Q_{II}^T(0,0) \end{array} \right) \qquad (34)$$

where $Q(0,0)$ fulfills

$$Q^T A_{00} Q = \Lambda_{00}, \quad A_{00} = \sum_{i=1}^{2} n_i A_i(0,0), \quad n_1 = -\frac{h_2}{h}, n_2 = -\frac{h_1}{h}, \quad h = \sqrt{h_1^2 + h_2^2}$$

The motive for defining $L(0,0)$ this way will be evident later. Furthermore, A_{00} is supposed to be non-singular. Let

$$
\begin{aligned}
L_0 &= \left(\begin{array}{ccc} L_{00} & \ldots & L_{\nu_1 0} \end{array} \right) & \in \mathbb{R}^{(\nu_1+1)d \times s_0}, \; s_0 = \sum_{i=0}^{\nu_1} d_1(i,0) \\
L_j &= \left(\begin{array}{cc} L_{0j} & L_{\nu_1 j} \end{array} \right) & \in \mathbb{R}^{(\nu_1+1)d \times s_j}, \; s_j = \sum_{i=0,\nu_1} d_1(i,j), \; j = 1, \ldots, \nu_2 - 1 \\
L_{\nu_2} &= \left(\begin{array}{ccc} L_{0\nu_2} & \ldots & L_{\nu_1 \nu_2} \end{array} \right) & \in \mathbb{R}^{(\nu_1+1)d \times s_{\nu_2}}, \; s_{\nu_2} = \sum_{i=0}^{\nu_1} d_1(i,\nu_2)
\end{aligned}
$$

25

The boundary conditions may thus be expressed as

$$L^T v = 0, \quad L = \begin{pmatrix} L_0 & & \\ & \ddots & \\ & & L_{\nu_2} \end{pmatrix} \in R^{(\nu_1+1)(\nu_2+1)d \times s}, \quad s = \sum_{j=0}^{\nu_2} s_j \qquad (35)$$

Obviously $\mathrm{rank}(L) = s$, i. e., L has full rank. Hence, the corresponding boundary projection is well-defined, and is given by

$$P = I - \Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T$$

where

$$\Sigma = \begin{pmatrix} \sigma_0 \Sigma_1 & & \\ & \ddots & \\ & & \sigma_{\nu_2} \Sigma_1 \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} \sigma_0 I & & \\ & \ddots & \\ & & \sigma_{\nu_1} I \end{pmatrix}, \quad I \in R^{d \times d}$$

It is possible to simplify the expression for P in this case. We have

$$\Sigma^{-1} L = \begin{pmatrix} \Sigma_1^{-1} L_0 / \sigma_0 & & \\ & \ddots & \\ & & \Sigma_1^{-1} L_{\nu_2} / \sigma_{\nu_2} \end{pmatrix}$$

But $\Sigma_1^{-1} L_j = L_j H_j$, where

$$H_j = \begin{pmatrix} I/\sigma_0 & \\ & I/\sigma_{\nu_1} \end{pmatrix} \in R^{s_j \times s_j} \quad j = 1, \ldots, \nu_2 - 1$$

$$H_j = \begin{pmatrix} I/\sigma_0 & & \\ & \ddots & \\ & & I/\sigma_{\nu_1} \end{pmatrix} \in R^{s_j \times s_j} \quad j = 0, \nu_2$$

Hence

$$\Sigma^{-1} L = LH, \quad H = \begin{pmatrix} H_0/\sigma_0 & & \\ & \ddots & \\ & & H_{\nu_2}/\sigma_{\nu_2} \end{pmatrix} \in R^{s \times s}$$

Clearly, H is invertible. We therefore arrive at

$$P = I - LH(L^T LH)^{-1} L^T = I - L(L^T L)^{-1} L^T$$

i. e., P is independent of Σ.

The semi-discrete system can now be defined as

$$v_t = P \left(\sum_{i=1}^{2} A_i D_i v + F \right) \qquad (36)$$
$$v(0) = f$$

It will next be shown that the solution to the system above satisfies an energy estimate.

26

Proposition 4.1 *Let $(\cdot,\cdot)_h$ be given by (5) and suppose that D_1 and D_2 satisfy the conclusion of proposition 2.3. If P is defined by (10) and (35), then the solution of (36) satisfies an energy estimate*

$$||v(t)||_h^2 + \int_0^t ||v(\tau)||_\Gamma^2 d\tau \leq Ke^{\alpha' t}\left(||f||_h^2 + \int_0^t ||F(\tau)||_h^2 d\tau\right)$$

where the boundary energy $||\cdot||_\Gamma$ is given by ($\nu_1 = \nu_2 = \nu$ for convenience)

$$||v(\tau)||_\Gamma^2 = h_2\sum_{j=0}^{\nu}\sigma_j\left(|v_{0j}|^2 + |v_{\nu j}|^2\right) + h_1\sum_{i=0}^{\nu}\sigma_i\left(|v_{i0}|^2 + |v_{i\nu}|^2\right)$$

Proof:
From propositions 2.5 and 2.4 we obtain

$$\frac{d}{dt}||v||_h^2 = 2\sum_{i=1}^{2}(v, A_iD_iv)_h + 2(v, F)_h$$

From proposition 2.3 and corollary 2.1 it follows that (v is only supported in a neighborhood of $(0,0)$)

$$(v, A_1D_1v)_h = -\frac{1}{2}\left(h_2\sum_{j=0}^{\nu}\sigma_j v_{0j}^T A_1 v_{0j} + (v, [D_1, A_1]v)_h\right)$$
$$(v, A_2D_2v)_h = -\frac{1}{2}\left(h_1\sum_{i=0}^{\nu}\sigma_i v_{i0}^T A_2 v_{i0} + (v, [D_2, A_2]v)_h\right)$$

Thus, by lemma 2.3 we have

$$\frac{d}{dt}||v||_h^2 \leq -h_1\sum_{i=0}^{\nu}\sigma_i v_{i0}^T A_2 v_{i0} - h_2\sum_{j=0}^{\nu}\sigma_j v_{0j}^T A_1 v_{0j} + \left(\sum_{i=1}^{2}\rho([D_i, A_i]) + 1\right)||v||_h^2 + ||F||_h^2$$

In the first sum the outward unit normal is $n = (0, -1)$, and in the second $n = (-1, 0)$. Except for the origin, the boundary terms are of exactly the same form as in the one-dimensional case. Eqs. (30), (31) thus imply that

$$-v_{i0}^T A_2 v_{i0} = \varphi_{i0}^T \Lambda_{i0}\varphi_{i0} \leq -\frac{\gamma_{i0}}{2}|\varphi_{0i}|^2 = -\frac{\gamma_{i0}}{2}|v_{0i}|^2, \quad i \geq 1$$

and a similar inequality holds for the other terms. At the origin we get

$$-h_2\sigma_0 v_{00}^T A_1 v_{00} - h_1\sigma_0 v_{00}^T A_2 v_{00} = h\sigma_0\varphi_{00}^T\Lambda_{00}\varphi_{00} \leq -h\sigma_0\frac{\gamma_{00}}{2}|\varphi_{00}|^2$$

But $h \geq (h_1 + h_2)/\sqrt{2}$. Hence

$$-h_2\sigma_0 v_{00}^T A_1 v_{00} - h_1\sigma_0 v_{00}^T A_2 v_{00} \leq -h_1\sigma_0\frac{\gamma_{00}}{2\sqrt{2}}|v_{00}|^2 - h_2\sigma_0\frac{\gamma_{00}}{2\sqrt{2}}|v_{00}|^2$$

Since $\Lambda(x)$ is uniformly non-singular it follows that $\gamma \equiv \inf(\gamma_{00}/\sqrt{2}, \gamma_{i0}, \gamma_{0j}) > 0$. Because of γ_{00}, the constant γ will in general be smaller than the corresponding constant of the analytic energy estimate. We thus arrive at

$$\frac{d}{dt}||v||_h^2 + \frac{\gamma}{2}||v||_\Gamma^2 \leq \left(\sum_{i=1}^{2} \rho([D_i, A_i]) + 1 \right) ||v||_h^2 + ||F||_h^2$$

which proves the proposition $(K = \max(1, 2/\gamma))$. $\qquad\qquad\qquad\qquad\Box$

4.2   The Heat Equation

The analysis of a homogeneous Dirichlet condition is straightforward, even if the domain of definition Ω is non-trivial. The problem lies in discretizing the Neumann conditions properly. This was clear in one space dimension. In two dimensions the occurrence of corners certainly complicates the analysis. To gain insight we shall begin by looking at a simple model problem.

The two-dimensional heat equation reads

$$\begin{aligned}
u_t &= u_{x_1 x_1} + u_{x_2 x_2} \quad & x \in \Omega = (0,1) \times (0,1) \\
u_n(x,t) &= 0 & x \in \Gamma \\
u(x,0) &= f(x)
\end{aligned}$$

where u_n is the normal derivative of u. Again, we focus our attention to a neighborhood of $(0,0)$. The boundary conditions are discretized as

$$\frac{1}{h_1} \sum_{k=0}^{r} d_{0k} v_{kj} = 0, \quad j = 0, \ldots, r \qquad \frac{1}{h_2} \sum_{k=0}^{r} d_{0k} v_{ik} = 0, \quad i = 0, \ldots, r \qquad (37)$$

or, equivalently,

$$(D_1 v)_{0j} = 0, \quad j = 0, \ldots, r \qquad (D_2 v)_{i0} = 0, \quad i = 0, \ldots, r \qquad (38)$$

where D_1 and D_2 are defined by proposition 2.3. The conditions above imply that two boundary conditions are prescribed at the origin for the discrete problem. This approach is natural from the intuitive point of view, in that gradients at the origin may be interpreted as one-sided limits from the interior. For the time being we ignore this technicality. It will later be shown how it can be overcome. When deriving the projection operator it is convenient to cast the boundary conditions into yet a another form. Define the boundary operators L_{1j} and L_{2i} through

$$L_{1j}^T v \equiv (D_1 v)_{0j} = 0, \quad j = 0, \ldots, r \qquad L_{2i}^T v \equiv (D_2 v)_{i0} = 0, \quad i = 0, \ldots, r \qquad (39)$$

where

$$L_{1j}^T = \left(\begin{array}{ccccccc} 0 & \ldots & 0 & \frac{1}{h_1} \sum_{k=0}^{r} d_{0k} e_k^T & 0 & \ldots & 0 \end{array} \right) \in \mathrm{R}^{1 \times (\nu_1 + 1)(\nu_2 + 1)}, j = 0, \ldots, r$$

$$L_{2i}^T = \left(\begin{array}{ccccc} \frac{d_{00}}{h_2} e_i^T & \ldots & \frac{d_{0r}}{h_2} e_i^T & 0 & \ldots & 0 \end{array} \right) \qquad \in \mathrm{R}^{1 \times (\nu_1 + 1)(\nu_2 + 1)}, i = 0, \ldots, r$$

Here $\{e_i\}$ is the canonical basis in R^{ν_1+1}. The boundary conditions can thus be written in standard form

$$L^T v = 0 \quad L = \begin{pmatrix} L_{10} & \ldots & L_{1r} & L_{20} & \ldots & L_{2r} \end{pmatrix} \in R^{2(r+1) \times (\nu_1+1)(\nu_2+1)} \qquad (40)$$

We know that the corresponding projection operator is well-defined iff $\mathrm{rank}(L) = 2(r+1)$.

Lemma 4.1 *The columns of L (40) are linearly dependent. In particular, $\mathrm{rank}(L) \leq 2r + 1$.*

Proof:
To investigate linear dependence we study

$$\sum_{j=0}^{r} \alpha_j L_{1j} + \sum_{j=0}^{r} \beta_j L_{2j} = 0$$

which is equivalent to

$$\sum_{k=0}^{r} \left(\alpha_j h_2 d_{0k} + \beta_k h_1 d_{0j} \right) e_k = 0, \quad j = 0, \ldots, r$$

Since $\{e_k\}$ is an ortho–normal system it follows that

$$\alpha_j h_2 d_{0k} + \beta_k h_1 d_{0j} = 0, \quad j, k = 0, \ldots, r$$

which obviously has the non-trivial solution

$$\alpha_j = d_{0j} \quad \beta_j = -\frac{h_2}{h_1} d_{0j}, \quad j = 0, \ldots, r$$

and the lemma is proved. $\qquad \square$

As a consequence of lemma 4.1 the projection formulation breaks down. If, however, we change the boundary condition at the origin to

$$L_{0\chi}^T v \equiv \left((1 - \chi) L_{10}^T + \chi L_{20}^T \right) v = 0 \quad 0 \leq \chi \leq 1 \qquad (41)$$

and leave the boundary conditions at the remaining points unchanged we get a well-defined projection operator, since

$$L = \begin{pmatrix} L_{11} & \ldots & L_{1r} & L_{0\chi} & L_{21} & \ldots & L_{2r} \end{pmatrix} \in R^{2r+1 \times (\nu_1+1)(\nu_2+1)} \qquad (42)$$

has full rank.

Lemma 4.2 *The columns of L (42) are linearly independent. In particular, $\mathrm{rank}(L) = 2r + 1$.*

29

Proof:

Again we study

$$\sum_{j=1}^{r} \alpha_j L_{1j} + \gamma L_{0\chi} + \sum_{j=1}^{r} \beta_j L_{2j} = 0$$

which is the same as

$$\frac{d_{00}}{h_2} \sum_{k=1}^{r} \beta_k e_k + \gamma \left((1-\chi) \frac{1}{h_1} \sum_{k=0}^{r} d_{0k} e_k + \chi \frac{d_{00}}{h_2} e_0 \right) = 0$$

$$\frac{d_{0j}}{h_2} \sum_{k=1}^{r} \beta_k e_k + \alpha_j \frac{1}{h_1} \sum_{k=0}^{r} d_{0k} e_k + \gamma \chi \frac{d_{0j}}{h_2} e_0 = 0 \quad j = 1, \ldots, r$$

The first component of the first equation yields $\gamma(h_2(1-\chi) + h_1\chi)d_{00} = 0$. Since $d_{00} \neq 0$ for any operator satisfying proposition 2.3, and since $h_i > 0$, $0 \leq \chi \leq 1$, necessarily $\gamma = 0$. From the remaining components of the first equation we then obtain $\beta_j = 0$, $j = 1, \ldots, r$, which in turn implies $\alpha_j = 0$, $j = 1, \ldots, r$. The columns of L are thus linearly independent, i. e., L has full rank. $\qquad\square$

Before proceeding with the energy estimate, one more technical lemma is needed. Let $L_{0\chi_1}$ and $L_{0\chi_2}$ be defined by (41), and let

$$L = \left(\begin{array}{ccccccccc} L_{11} & \ldots & L_{1r} & L_{0\chi_1} & L_{0\chi_2} & L_{21} & \ldots & L_{2r} \end{array} \right) \in \mathbb{R}^{2(r+1)\times(\nu_1+1)(\nu_2+1)} \qquad (43)$$

Lemma 4.3 *The columns of L (43) are linearly dependent. In particular,* $\mathrm{rank}(L) \leq 2r + 1$.

Proof:

Consider

$$\sum_{j=1}^{r} \alpha_j L_{1j} + \gamma_1 L_{0\chi_1} + \gamma_2 L_{0\chi_2} + \sum_{j=1}^{r} \beta_j L_{2j} = 0$$

Obviously the lemma is true for $\chi_1 = \chi_2$. In the following we thus assume $\chi_1 \neq \chi_2$. The equation above can be rewritten as

$$\sum_{j=0}^{r} \alpha_j L_{1j} + \sum_{j=0}^{r} \beta_j L_{2j} = 0 \qquad (44)$$

where

$$\left(\begin{array}{cc} 1-\chi_1 & 1-\chi_2 \\ \chi_1 & \chi_2 \end{array} \right) \left(\begin{array}{c} \gamma_1 \\ \gamma_2 \end{array} \right) = \left(\begin{array}{c} \alpha_0 \\ \beta_0 \end{array} \right)$$

According to lemma 4.1, eq. (44) has the non-trivial solution

$$\alpha_j = d_{0j} \qquad \beta_j = -\frac{h_2}{h_1} d_{0j}, \quad j = 0, \ldots, r$$

30

whence

$$\gamma_1 = d_{00}\left(\chi_2 + \frac{h_2}{h_1}(1-\chi_2)\right)/(\chi_2-\chi_1)$$
$$\gamma_2 = -d_{00}\left(\chi_1 + \frac{h_2}{h_1}(1-\chi_1)\right)/(\chi_2-\chi_1)$$
$$\alpha_j = d_{0j} \quad \beta_j = -\frac{h_2}{h_1}d_{0j}, \quad j = 1,\ldots,r$$

solves the original equation. The lemma is proved. □

Proposition 4.2 *Let P be given by proposition 2.5, where L is defined by (42). Then $L_{10}^T P = L_{20}^T P = 0$.*

Proof:
Clearly, $L^T P = 0$. Furthermore, $L_{10}, L_{20} = L_{0\chi}$ for $\chi = 0, 1$, respectively. But then, by lemma 4.3,

$$L_{10} = L\alpha_1 \qquad L_{20} = L\alpha_2$$

for some vectors $\alpha_1, \alpha_2 \in \mathrm{R}^{2r+1}$. This proves the proposition. □

Remark: Suppose that v is a vector such that $v = Pv$, where P is as in the previous proposition. Then $L_{10}v = L_{20}v = 0$, i. e., $(D_1v)_{00} = (D_2v)_{00} = 0$. In other words, by requiring that the boundary condition at the origin hold for a specific convex combination we actually get the stronger result $(D_1v)_{00} = (D_2v)_{00} = 0$. Thus, we need not overspecify at the corners, cf. eq. (37). In the appendix we give a direct proof that $L_{10}^T P = 0$ for $L_{0\chi}$ with $\chi = 0.5$.

The semi-discrete heat equation is given by

$$\begin{aligned} v_t &= P\left(D_1^2 + D_2^2\right)v \\ v(0) &= f \end{aligned} \tag{45}$$

Proposition 4.3 *Let $(\cdot,\cdot)_h$ be given by (5) and suppose that D_1 and D_2 satisfy the conclusion of proposition 2.3. If P is defined by (10) and (42), then the solution of (45) satisfies an energy estimate*

$$\|v(t)\|_h \le \|f\|_h$$

Proof:
The energy method gives

$$\frac{d}{dt}\|v\|_h^2 = 2(v, D_1^2 v)_h + 2(v, D_2^2 v)_h$$

By proposition 2.3 (v is supported only in a neighborhood of the origin),

$$(v, D_1^2 v)_h = -h_2 \sum_{j=0}^{r} \sigma_j v_{0j}(D_1v)_{0j} - \|D_1v\|_h^2$$

According to propositions 2.4, 2.5 and 4.2 we have

$$(D_1v)_{0j} = L_{1j}^T v = 0, \quad j = 0,\ldots,r$$

The remaining term $(v, D_2^2 v)_h$ is treated similarly, and the proposition follows. □

31

4.3 Parabolic Systems

Consider

$$u_t = \sum_{i,j=1}^{2} A_{ij} u_{x_i x_j} + F, \qquad x \in \Omega = (0,1) \times (0,1), \quad u \in \mathbb{R}^d$$
$$u(x,0) = f(x), \qquad\qquad x = (x_1, x_2) \tag{46}$$
$$L_0(x)u(x,t) + L_1(x)u_n(x,t) = 0, \quad x \in \Gamma$$

The assumptions on L_0, L_1 in (17) are supposed to hold pointwise for each $x \in \Gamma$. Furthermore, we require that assumption 3.1 with $A = A_{ii}$ be valid on $x_i = 0$, $i = 1,2$. In particular, the conclusion of lemma 3.2 holds for each boundary point. It will be assumed that (46) is strongly parabolic, i. e., there are vectors $u_i(x,t) \in \mathbb{R}^d$, $i = 1,2$, such that

$$\sum_{i,j=1}^{2} u_i(x,t)^T A_{ij}(x,t) u_j(x,t) \geq 2\delta \sum_{i=1}^{2} |u_i(x,t)|^2$$

for all $x \in \Omega$, $t \geq 0$. If the matrices $A_{ij} \neq 0$, $i \neq j$, then the assumptions must be strengthened. The energy method applied to one of the cross terms yields (u is supported only at the origin, $A_{12} = $ const for simplicity, Ω is the unit square)

$$(u, A_{12} u_{x_1 x_2}) = - \int_{x_1=0} u^T A_{12} u_{x_2} dx_2 - (u_{x_1}, A_{12} u_{x_1})$$

In general we cannot get an estimate of $u_{x_2}(0, x_2, t)$ in the boundary integral. It is therefore natural to require

Assumption 4.1 $A_{ij}^T = A_{ij}$, $i \neq j$.

Remark: Neglecting scaling factors we have

$$A_{12} = A_{21} = \frac{1}{\rho} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

for the Navier-Stokes equations (ρ denotes the density). Clearly, assumption 4.1 is fulfilled.

If assumption 4.1 holds one can integrate by parts once more to obtain

$$(u, A_{12} u_{x_1 x_2}) = \frac{1}{2} u^T A_{12}(0,0,t) u - (u_{x_1}, A_{12} u_{x_2})$$

In two dimensions we cannot eliminate the boundary terms by means of Sobolev inequalities, since they would involve L^2-norms of $u_{x_1 x_1}$ and so forth. This motivates

Assumption 4.2 *Let $u(x,t)$ satisfy*

$$L_0(0,0)u + L_1(0,0)u_n = 0$$

at the origin. Then

$$u^T A_{ij}(0,0,t)u = 0, \quad i \neq j$$

Remark: This assumption ensures an energy estimate for the continuous problem in case of a non-smooth boundary, and couples the cross terms of the differential operator to the boundary conditions at the origin. In case of the Navier-Stokes equations one has zero velocity at the origin. Hence, the state vector becomes $u^T = (\begin{array}{cccc} \rho & 0 & 0 & p \end{array})$, which implies assumption 4.2.

The discrete boundary conditions are formulated as (D_1 and D_2 are defined by proposition 2.3)

$$L_{1j}^T v \equiv L_0(0, jh_2)v_{0j} + L_1(0, jh_2)(D_1 v)_{0j} = 0, \quad j = 0, \ldots, r$$

$$L_{2i}^T v \equiv L_0(ih_1, 0)v_{i0} + L_1(ih_1, 0)(D_2 v)_{i0} = 0, \quad i = 0, \ldots, r \tag{47}$$

where

$$L_{1j}^T = \left(\begin{array}{ccccccc} 0 & \ldots & 0 & L_0(0, jh_2)e_0^T + L_1(0, jh_2)\dfrac{1}{h_1}\sum_{k=0}^{r} d_{0k}e_k^T & 0 & \ldots & 0 \end{array} \right)$$

$$L_{2i}^T = \left(\begin{array}{cccccc} \left(L_0(ih_1, 0) + L_1(ih_1, 0)\dfrac{d_{00}}{h_2}\right)e_i^T & \ldots & L_1(ih_1, 0)\dfrac{d_{0r}}{h_2}e_i^T & 0 & \ldots & 0 \end{array} \right)$$

and $e_i^T = (\begin{array}{ccccccc} 0 & \ldots & 0 & I & 0 & \ldots & 0 \end{array}) \in \mathrm{R}^{d \times (\nu_1 + 1)d}$. The boundary conditions can be expressed in the usual form as

$$L^T v = 0 \quad L = \left(\begin{array}{ccccccc} L_{11} & \ldots & L_{1r} & L_{0\chi} & L_{21} & \ldots & L_{2r} \end{array} \right) \in \mathrm{R}^{(2r+1)d \times (\nu_1 + 1)(\nu_2 + 1)d} \tag{48}$$

where

$$L_{0\chi} \equiv (1 - \chi)L_{10} + \chi L_{20} \quad 0 \leq \chi \leq 1$$

Lemma 4.4 *The columns of L (48) are linearly independent for sufficiently small step lengths h_1 and h_2. In particular,* $\mathrm{rank}(L) = (2r + 1)d$.

Proof:
Imitating the proof of lemma 4.2 gives

$$\gamma \left[L_0(0,0) + d_{00} \left(\frac{1 - \chi}{h_1} + \frac{\chi}{h_2} \right) L_1(0,0) \right] = 0$$

33

By lemma 3.2 the expression inside the brackets is non-singular for h_1, h_2 sufficiently small. Hence, $\gamma = 0$, which in turn implies $\alpha_j = \beta_j = 0$, $j = 1, \ldots, r$. Since the columns of each block L_{1j}, L_{2j} and $L_{0\chi}$ are linearly independent, the lemma follows. $\square$

The semi-discrete parabolic system reads

$$v_t = P\left(\sum_{i,j=1}^{2} A_{ij} D_i D_j v + F\right),$$
$$v(0) = f \tag{49}$$

where P is defined by proposition 2.5 and by (48). Unfortunately, assumption 4.2 is not sufficient for the semi-discrete problem. We need

Assumption 4.3 *Let v satisfy*

$$L_0(0,0)v_{00} + L_1(0,0)\left((1-\chi)(D_1 v)_{00} + \chi(D_2 v)_{00}\right) = 0 \qquad 0 \le \chi \le 1$$

at the origin. Then

(i)

$$v_{00}^T A_{ij}(0,0,t)v_{00} = 0, \quad i \ne j$$

(ii)

$$v_{00}^T A_{11}(0,0,t) = v_{00}^T A_{22}(0,0,t)$$

Remark: The first requirement is identical to that of assumption 4.2. The second, however, appears only in the discrete case. We note that assumption 4.3 holds for the Navier-Stokes equations, since A_{11} and A_{22} are given by

$$A_{11} = \frac{1}{\rho}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & c_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -c_2 p/\rho & 0 & 0 & c_2 \end{pmatrix} \qquad A_{22} = \frac{1}{\rho}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_1 & 0 \\ -c_2 p/\rho & 0 & 0 & c_2 \end{pmatrix}$$

Hence, $v_{00}^T A_{11} = v_{00}^T A_{22} = c_2(\; -p^2/\rho^2 \quad 0 \quad 0 \quad p/\rho\;)$.

Proposition 4.4 *Let $(\cdot,\cdot)_h$ be given by (5) and suppose that D_1 and D_2 satisfy the conclusion of proposition 2.3. If P is defined by (10) and (48), and if assumptions 4.1 and 4.3 hold, then the solution of (49) satisfies an energy estimate*

$$||v(t)||_h^2 + \int_0^t ||v(\tau)||_\Gamma^2 d\tau \le e^{(\alpha' + \mathcal{O}(h))t}\left(||f||_h^2 + \int_0^t ||F(\tau)||_h^2 d\tau\right)$$

34

Proof:

The energy method gives

$$\frac{d}{dt}||v||_h^2 \leq -2\sum_{j=1}^{2}\left(h_2\sum_{k=0}^{r}\sigma_k v_{0k}^T A_{1j}(D_j v)_{0k} + h_1\sum_{k=0}^{r}\sigma_k v_{k0}^T A_{2j}(D_j v)_{k0}\right)$$
$$-2\delta\sum_{i=1}^{2}||D_i v||_h^2 + (K_0 + \mathcal{O}(h))\,||v||_h^2 + ||F||_h^2$$

where K_0 depends on $||[D_i, A_{ii}]||_h$, $i = 1, 2$ and $\rho([D_i, A_{ij}])$, $i \neq j$. The first cross term can be written as (v has compact support).

$$-h_2\sum_{k=0}^{r}\sigma_k v_{0k}^T A_{12}(D_2 v)_{0k} = -h_2\sum_{k=0}^{\nu}\sigma_k v_{0k}^T A_{12}\left(\frac{1}{h_2}\sum_{l=0}^{\nu}d_{kl}v_{0l}\right) \equiv -(v_0, \tilde{A}_{12}\tilde{D}_2 v_0)_{h_2}$$

where $v_0^T = (v_{00}^T \ldots v_{0\nu}^T)$, and where $\tilde{D}_2$ satisfies (1) with respect to the one-dimensional scalar product $(\cdot, \cdot)_{h_2}$. Hence,

$$-(v_0, \tilde{A}_{12}\tilde{D}_2 v_0)_{h_2} = \frac{1}{2}v_{00}^T A_{12}(0, 0, t)v_{00} + \frac{1}{2}(v_0, [\tilde{D}_2, \tilde{A}_{12}]v_0)_{h_2}$$

By assumption 4.3 the boundary terms vanish. The remaining cross term is treated in a similar vein.

Next, we take care of the boundary terms corresponding to the pure second differences. Only the origin needs to be analyzed, since the other boundary points are treated exactly as in the proof of proposition 3.2. At the origin we get

$$-h_2\sigma_0 v_{00}^T A_{11}(D_1 v)_{00} - h_1\sigma_0 v_{00}^T A_{22}(D_2 v)_{00} =$$
$$-(h_1 + h_2)\sigma_0 v_{00}^T\left((1 - \chi)A_{11}(D_1 v)_{00} + \chi A_{22}(D_2 v)_{00}\right), \quad \chi = \frac{h_1}{h_1 + h_2}$$

and, by assumption 4.3,

$$-h_2\sigma_0 v_{00}^T A_{11}(D_1 v)_{00} - h_1\sigma_0 v_{00}^T A_{22}(D_2 v)_{00} =$$

$$-(h_1 + h_2)\sigma_0 v_{00}^T A_{11}\left((1 - \chi)(D_1 v)_{00} + \chi(D_2 v)_{00}\right)$$

But $v = Pv$ implies $L_{0\chi}v = 0$, i. e., by (47)

$$L_0(0, 0)v_{00} + L_1(0, 0)\left((1 - \chi)(D_1 v)_{00} + \chi(D_2 v)_{00}\right) = 0$$

In particular, $L_0^{II} v_{00} = 0$. Partition $v = v' + v''$ where $v'_{ij} \in \ker(L_1^I)$, $v''_{ij} \in \ker(L_1^I)^\perp$. Assumption 3.1 then gives

$$-h_2\sigma_0 v_{00}^T A_{11}(D_1 v)_{00} - h_1\sigma_0 v_{00}^T A_{22}(D_2 v)_{00} =$$

$$-(h_1 + h_2)\sigma_0 v_{00}^T A_{11}\left((1 - \chi)(D_1 v'')_{00} + \chi(D_2 v'')_{00}\right)$$

35

By construction

$$L_0(0,0)v_{00} + L_1(0,0)\left((1-\chi)(D_1v'')_{00} + \chi(D_2v'')_{00}\right) = 0$$

which can be solved in exactly the same way as the corresponding equation in the proof of proposition 3.2. Hence,

$$-h_2\sigma_0 v_{00}^T A_{11}(D_1v)_{00} - h_1\sigma_0 v_{00}^T A_{22}(D_2v)_{00} = h_2\sigma_0 v_{00}^T A_{11}\tilde{L}_1^{-1}L_0v_{00} + h_1\sigma_0 v_{00}^T A_{22}\tilde{L}_1^{-1}L_0v_{00}$$

where we again have invoked assumption 4.3. We thus arrive at

$$
\begin{aligned}
\frac{d}{dt}||v||_h^2 + ||v||_\Gamma^2 \leq\ & \left(2|A_{11}\tilde{L}_1^{-1}L_0|_{2,\infty} + \rho([\tilde{D}_2,\tilde{A}_{12}]) + 1\right) h_2 \sum_{k=0}^r \sigma_k|v_{0k}|^2 \\
& + \left(2|A_{22}\tilde{L}_1^{-1}L_0|_{1,\infty} + \rho([\tilde{D}_1,\tilde{A}_{21}]) + 1\right) h_1 \sum_{k=0}^r \sigma_k|v_{k0}|^2 \\
& - 2\delta\sum_{i=1}^2 ||D_iv||_h^2 + (K_0 + \mathcal{O}(h))\,||v||_h^2 + ||F||_h^2
\end{aligned}
$$

where $|v|_{1,\infty} = \sup(|v_{k0}|)$ and $|v|_{2,\infty} = \sup(|v_{0k}|)$. Replacing $\rho([\tilde{D}_i,\tilde{A}_{ji}])$ by $|A'_{ji}|_{i,\infty}$, $i \neq j$, one obtains the coefficients of the boundary terms of the analytic energy estimate. They are thus identical if the coefficient matrices are constant or if we use the standard second order method. Finally, the boundary terms of the right hand are eliminated by applying the one-dimensional Sobolev inequality 2.6 in the x_1- and x_2-directions, respectively. This proves the proposition. □

Remark: It is clear from the proof that (49) is strictly stable if the coefficient matrices are constant, or if $A_{ii}^T = A_{ii}$, $i = 1,2$ and the second-order method (3) is used.

4.4 Hyperbolic-Parabolic Systems

In this section we merely formulate the problem and state the main result. The reader is asked to fill in the details. We consider the mixed hyperbolic-parabolic system

$$
\begin{aligned}
u_t &= \sum_{i,j=1}^2 A_{ij}u_{x_ix_j} + \sum_{i=1}^2 C_{1i}v_{x_i} + F & x \in \Omega \\
v_t &= \sum_{i=1}^2 B_iv_{x_i} + \sum_{i=1}^2 C_{2i}u_{x_i} + G & u \in \mathrm{R}^{d_1}, v \in \mathrm{R}^{d_2}
\end{aligned}
$$

$$M(x) = \begin{pmatrix} M^I(x) \\ 0 \end{pmatrix}$$

$$
\begin{aligned}
L_1(x)u_n(x,t) + L_0(x)u(x,t) + M(x)v(x,t) &= 0 & x \in \Gamma \\
\varphi_I(x,t) &= S(x)\varphi_{II}(x,t) + R(x)u(x,t) & x \in \Gamma
\end{aligned}
$$

$$
\begin{aligned}
u(x,0) &= f(x) \\
v(x,0) &= \phi(x)
\end{aligned}
$$

$$(50)$$

The structural hypotheses on the parabolic and hyperbolic parts of the principal operator are identical to those in sections 4.3 and 4.1. This remark also pertains to the boundary conditions. In particular, the characteristic variables φ_I and φ_{II} are defined as in section 4.1. The principal operator is described by the A_{ij}'s and B_i's; the lower order coupling is determined by the C_{ij}'s. Similarly, in the boundary conditions the coupling is expressed by M and R.

Define $w^T = (\ u^T \quad v^T\ )$ and let P be the projection corresponding to the boundary conditions of (50). The semi-discrete system is then defined as

$$
w_t = P\left(\sum_{i,j=1}^{2} \tilde{A}_{ij}\tilde{D}_i\tilde{D}_j w + \sum_{i=1}^{2} \tilde{B}_i\tilde{D}_i w + \sum_{i=1}^{2} \tilde{C}_i\tilde{D}_i w + \tilde{F}\right) \qquad w_{ij} = \begin{pmatrix} u_{ij} \\ v_{ij} \end{pmatrix} \qquad (51)
$$
$$
w(0) = \psi
$$

where

$$
\tilde{A}_{ij} = \begin{pmatrix} \begin{pmatrix} A_{ij}(0,0,t) & 0 \\ 0 & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} A_{ij}(1,1,t) & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix}
$$

$$
\tilde{B}_i = \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & B_i(0,0,t) \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & 0 \\ 0 & B_i(1,1,t) \end{pmatrix} \end{pmatrix}
$$

$$
\tilde{C}_i = \begin{pmatrix} \begin{pmatrix} 0 & C_{1i}(0,0,t) \\ C_{2i}(0,0,t) & 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} 0 & C_{1i}(1,1,t) \\ C_{2i}(1,1,t) & 0 \end{pmatrix} \end{pmatrix}
$$

The forcing function $\tilde{F}$ and the initial data ψ are defined in a similar fashion.

Proposition 4.5 *Let* $(\cdot,\cdot)_h$ *be given by (5) and suppose that* $\tilde{D}$ *satisfies the conclusion of proposition 2.1. Then the solution of (51) satisfies an energy estimate*

$$
\|u(t)\|_h^2 + \|v(t)\|_h^2 + \int_0^t \left(\|u(\tau)\|_\Gamma^2 + \|v(\tau)\|_\Gamma^2\right) d\tau
$$
$$
\leq K e^{(\alpha'+\mathcal{O}(h))t}\left(\|f\|_h^2 + \|\phi\|_h^2 + \int_0^t \left(\|F(\tau)\|_h^2 + \|G(\tau)\|_h^2\right) d\tau\right)
$$

5   General Domains and Strict Stability

Nothing has been said about strict stability in two dimensions thus far. The purpose of strict stability is to ensure the same growth rate of the discrete and analytic solutions. If the analytic problem is defined on a curvilinear domain Ω, then there must exist a diffeomorphism $\xi = \xi(x)$ of Ω onto the unit square $(0,1) \times (0,1)$ in order for the finite difference method to be well-defined. Consequently, a constant coefficient problem in the original domain may be transformed to a variable coefficient problem on the unit square, which may account for a non-physical growth in the discrete estimate.

Let $\xi = \xi(x)$ be a diffeomorphism of Ω onto $I = (0,1) \times (0,1)$. The following identities are readily established

$$
\frac{\partial x_1}{\partial \xi_1} = J^{-1}\frac{\partial \xi_2}{\partial x_2} \qquad \frac{\partial x_2}{\partial \xi_1} = -J^{-1}\frac{\partial \xi_2}{\partial x_1}
$$
$$
\frac{\partial x_1}{\partial \xi_2} = -J^{-1}\frac{\partial \xi_1}{\partial x_2} \qquad \frac{\partial x_2}{\partial \xi_2} = J^{-1}\frac{\partial \xi_1}{\partial x_1}
\qquad\qquad J = \det\left(\frac{\partial \xi}{\partial x}\right) \tag{52}
$$

which in turn implies

$$
\sum_{i=1}^{2} \left(J^{-1}\partial \xi_i\right)_{\xi_i} = 0 \tag{53}
$$

where ∂ denotes the two-dimensional gradient operator. We require that $\xi(x)$ be uniformly non-singular, i. e., there exists a constant $\delta > 0$ such that $J^{-1} \geq \delta$ on Ω, For later use we record the normal and tangential derivatives u_{n_i} and u_{τ_i} at the boundaries corresponding to $\xi_i = 0$, $i = 1,2$:

$$
u_{\tau_i} = (-1)^i u_{\xi_j}/|x_{\xi_j}|
$$
$$
\qquad\qquad i = 1,2 \;\; j \neq i \tag{54}
$$
$$
u_{n_i} = -(\partial \xi_i \cdot \partial \xi_i u_{\xi_i} + \partial \xi_i \cdot \partial \xi_j u_{\xi_j})/|\partial \xi_i|
$$

where the boundary Γ of the domain Ω has been parametrized in the positive direction.

The analytic scalar product obeys

$$
(u,v) = \int_\Omega u^T(x)v(x)dr = \int_I u^T(x(\xi))v(x(\xi))J^{-1}d\xi
$$

which suggests the following semi-discrete scalar product

$$
\langle u,v\rangle_h = (u,J^{-1}v)_h = (J^{-1}u,v)_h \tag{55}
$$

where the last equality follows since J^{-1} and Σ are diagonal. Thus, each grid point is scaled with the cell volume. Similarly, the analytic boundary integrals can be parametrized as

$$
\int_{\Gamma_1} u^T(x)v(x)ds = \int_0^1 u^T(x(\xi_1,0))v(x(\xi_1,0))|x_{\xi_1}(\xi_1,0)|d\xi_1
$$

Hence, it is natural to define the boundary scalar product as

$$\langle u, v \rangle_\Gamma = \sum_{j=0}^{\nu} \sigma_j \left(s_{0j} u_{0j}^T v_{0j} + s_{\nu j} u_{\nu j}^T v_{\nu j} \right) + \sum_{i=0}^{\nu} \sigma_i \left(s_{i0} u_{i0}^T v_{i0} + s_{i\nu} u_{i\nu}^T v_{i\nu} \right) \tag{56}$$

where the arc lengths are defined as

$$s_{0j} = |x_{\xi_2}(0, jh_2)| h_2 \quad s_{i0} = |x_{\xi_1}(ih_1, 0)| h_1 \quad h_1 = \Delta \xi_1, \quad h_2 = \Delta \xi_2$$

with similar definitions for $s_{\nu j}$ and $s_{i\nu}$.

In order to prove stability we must have $Pv = v$. Since v will be the solution of equations like (60), proposition 2.4 implies that $PJ^{-1}v = J^{-1}v$. Therefore, it is natural to require

$$PJ^{-1} = J^{-1}P \tag{57}$$

For a general P this identity expresses a compatibility condition between the analytic boundary conditions and the mapping $\xi(x)$. Let P be given by (10) and (42). Then (57) certainly holds if

$$\begin{aligned} J_{ij} \equiv J(x(ih_1, jh_2)) = J_{i0}, \quad j = 0, \ldots, r \\ J_{ij} \equiv J(x(ih_1, jh_2)) = J_{0j}, \quad i = 0, \ldots, r \end{aligned} \tag{58}$$

which states that the mapping $\xi(x)$ is locally *isochoric* in the x_{ξ_i}-direction at the boundary, where $\partial / \partial \xi_i$ is the non-tangential derivative. In case of characteristic boundary conditions and Dirichlet conditions we have $r = 0$, and (58) is trivially satisfied. For general boundary conditions, however, (58) couples the boundary operator to the grid transformation (cf. assumption 4.3, which links the differential operator to the boundary operator).

5.1 Symmetric Hyperbolic Systems

Using (53) we recast (29) into a form that eliminates the need for the commutator in the semi-discrete case.

$$\left(J^{-1} u \right)_t = \frac{1}{2} \sum_{i=1}^{2} \left(\left(J^{-1} B_i u \right)_{\xi_i} + J^{-1} B_i u_{\xi_i} \right) - \frac{1}{2} J^{-1} \mathrm{div}(A) u + J^{-1} F \tag{59}$$

where

$$B_i = \partial \xi_i \cdot A = \frac{\partial \xi_i}{\partial x_1} A_1 + \frac{\partial \xi_i}{\partial x_2} A_2 \qquad \mathrm{div}(A) = \partial \cdot A = \frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2}$$

The boundary conditions are as in (29), except at the origin, where we require that the characteristic boundary conditions $\varphi_I(0, t) = S(0) \varphi_{II}(0, t)$ be satisfied for $\varphi(0, t) = Q_i^T(0) u(0, t)$, $i = 1, 2$, where $Q_i^T(0) A_i(0, t) Q_i(0)$ are diagonal, which means that two boundary conditions are prescribed at the corner. This situation occurs for the Euler equations at corners, where it is natural to require that both normal components of the velocity field be zero. Furthermore, this assumption simplifies the computations that

39

follow. The projection operator is still well-defined. It should be pointed out that the boundary condition at the origin is only used for the semi-discrete system.

Eq. (59) is discretized as

$$\left(J^{-1}v\right)_t = P\left(\frac{1}{2}\sum_{i=1}^{2}\left(D_iJ^{-1}B_iv + J^{-1}B_iD_iv\right) - \frac{1}{2}J^{-1}Cv + J^{-1}F\right) \qquad (60)$$

with

$$C = \begin{pmatrix} \text{div}(A)(x(0,0),t) & & \\ & \ddots & \\ & & \text{div}(A)(x(1,1),t) \end{pmatrix}$$

Proposition 5.1 *The approximation (60) is strictly stable.*

Proof:
The energy method yields (using $P(J^{-1}v) = J^{-1}v$, $PJ^{-1} = J^{-1}P \implies Pv = v$)

$$\frac{d}{dt}\langle v,v\rangle_h = \sum_{i=1}^{2}\left((v,D_iJ^{-1}B_iv)_h + (v,J^{-1}B_iD_iv)_h\right) - (v,J^{-1}Cv)_h + 2(v,J^{-1}F)_h$$

But (using $B_1(0,jh_2)$ instead of $B_1(x(0,jh_2),t)$ and so forth to make the notation less cumbersome)

$$(v,D_1J^{-1}B_1v)_h = -h_2\sum_{j=0}^{\nu}\sigma_j v_{0j}^T J^{-1}(0,jh_2)B_1(0,jh_2)v_{0j} - (D_1v,J^{-1}B_1v)_h$$

Since diagonal scalar products are used, assumption 2.1 holds *a fortiori*. Hence, ($B_1^T = B_1$)

$$(D_1v,J^{-1}B_1v)_h = (B_1J^{-1}D_1v,v)_h = (J^{-1}B_1D_1v,v)_h$$

where the last equality follows since B_1 and J^{-1} commute. Thus,

$$(v,D_1J^{-1}B_1v)_h = -h_2\sum_{j=0}^{\nu}\sigma_j v_{0j}^T J^{-1}(0,jh_2)B_1(0,jh_2)v_{0j} - (J^{-1}B_1D_1v,v)_h$$

with a similar relation for $(v,D_2J^{-1}B_2v)_h$. We thus arrive at

$$\frac{d}{dt}\langle v,v\rangle_h \leq -h_2\sum_{j=0}^{\nu}\sigma_j v_{0j}^T J^{-1}(0,jh_2)B_1(0,jh_2)v_{0j}$$

$$-h_1\sum_{i=0}^{\nu}\sigma_i v_{i0}^T J^{-1}(ih_1,0)B_2(ih_1,0)v_{i0}$$

$$+(|\text{div}(A)|_\infty + 1)\langle v,v\rangle_h + \langle F,F\rangle_h$$

40

By means of (52) it follows that

$$J^{-1}B_1 = \frac{\partial x_2}{\partial \xi_2}A_1 - \frac{\partial x_1}{\partial \xi_2}A_2$$

Apparently, x_{ξ_2} is a tangent vector of the curve $x(0, \xi_2)$. Hence

$$\left(\begin{array}{cc} -\dfrac{\partial x_2}{\partial \xi_2} & \dfrac{\partial x_1}{\partial \xi_2} \end{array} \right)$$

is an outward normal to $x(0, \xi_2) \in \Gamma$. The unit normal is then defined as

$$\left(\begin{array}{cc} n_1 & n_2 \end{array} \right) = \left(\begin{array}{cc} -\dfrac{\partial x_2}{\partial \xi_2} & \dfrac{\partial x_1}{\partial \xi_2} \end{array} \right) / |x_{\xi_2}|$$

Using the definition of the arc length we then obtain

$$-h_2 \sum_{j=0}^{\nu} \sigma_j v_{0j}^T J^{-1}(0, jh_2) B_1(0, jh_2) v_{0j} = \sum_{j=0}^{\nu} \sigma_j s_{0j} v_{0j}^T \left(n_1 A_1 + n_2 A_2 \right) v_{0j}$$

The boundary conditions are satisfied, whence

$$v_{0j}^T \left(n_1 A_1 + n_2 A_2 \right) v_{0j} \leq -\gamma_j |v_{0j}|^2$$

Letting $\gamma \equiv \inf(\gamma_j) > 0$ implies

$$-h_2 \sum_{j=0}^{\nu} \sigma_j v_{0j}^T J^{-1}(0, jh_2) B_1(0, jh_2) v_{0j} \leq -\gamma \sum_{j=0}^{\nu} \sigma_j s_{0j} |v_{0j}|^2$$

We have thus established

$$\frac{d}{dt} \langle v, v \rangle_h + \gamma \langle v, v \rangle_\Gamma \leq (|\mathrm{div}(A)|_\infty + 1) \langle v, v \rangle_h + \langle F, F \rangle_h$$

This is exactly the same estimate that one would get in the analytic case, and the proposition has been proved. □

5.2 The Heat Equation

The heat equation in self-adjoint form reads

$$\left(J^{-1} u \right)_t = \sum_{i,k=1}^{2} \left(J^{-1} \frac{\partial \xi_k}{\partial x_i} \left(\sum_{j=1}^{2} \frac{\partial \xi_j}{\partial x_i} u_{\xi_j} \right) \right)_{\xi_k} \tag{61}$$

At the boundary the normal derivative is set to zero. Define

$$M_{ij} = \begin{pmatrix} \dfrac{\partial \xi_j}{\partial x_i}(x(0,0)) & & \\ & \ddots & \\ & & \dfrac{\partial \xi_j}{\partial x_i}(x(1,1)) \end{pmatrix}$$

and

$$\tilde{D}_i = \sum_{j=1}^{2} M_{ij} D_j$$

Clearly, $\tilde{D}_i$ is a consistent approximation of $\partial/\partial x_i$. Let

$$\begin{aligned} L_{0n_1} &= -|\partial \xi_1| L_{10} - \frac{\partial \xi_1 \cdot \partial \xi_2}{|\partial \xi_1|} L_{20} \\ L_{0n_2} &= -|\partial \xi_2| L_{20} - \frac{\partial \xi_1 \cdot \partial \xi_2}{|\partial \xi_2|} L_{10} \end{aligned} \qquad (62)$$

be approximations of the normal derivatives at the origin. The boundary projection P is defined by (10) and (42) with $L_{0\chi}$ replaced by either of L_{0n_1} and L_{0n_2}.

It should be noted that P may no longer be unconditionally well-defined. Arguing exactly as in the proof of lemma 4.2 one obtains (using $L_{0\chi} = L_{0n_1}$)

$$- \left(|\partial \xi_1|^2 h_2 + \partial \xi_1 \cdot \partial \xi_2 h_1 \right) \gamma = 0$$

If $\partial \xi_1 \cdot \partial \xi_2 < 0$, i. e., at acute corners there is a possibility of a non-zero γ if

$$h_2 = -\frac{\partial \xi_1 \cdot \partial \xi_2}{|\partial \xi_1|^2} h_1 \qquad (63)$$

Hence, at acute corners we assume that h_1 and h_2 be such that (63) does not hold. Furthermore, lemma 4.3 is valid for $L_{0\chi_1} = L_{0n_1}$, $L_{0\chi_2} = L_{0n_2}$. This is obvious if $\partial \xi_1 \cdot \partial \xi_2 = \pm |\partial \xi_1||\partial \xi_2|$, because then $L_{0n_1} = \pm L_{0n_2}$. Otherwise, we obtain (44) where

$$- \begin{pmatrix} |\partial \xi_1| & \dfrac{\partial \xi_1 \cdot \partial \xi_2}{|\partial \xi_2|} \\ \dfrac{\partial \xi_1 \cdot \partial \xi_2}{|\partial \xi_1|} & |\partial \xi_2| \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}$$

which has a unique solution iff $|\partial \xi_1 \cdot \partial \xi_2| < |\partial \xi_1||\partial \xi_2|$.

The semi-discrete heat equation is now defined as

$$\left(J^{-1} v \right)_t = P \sum_{i,k=1}^{2} D_k \left(J^{-1} M_{ik} \left(\sum_{j=1}^{2} M_{ij} D_j v \right) \right) \qquad (64)$$

42

Proposition 5.2 *Assume that the mapping $\xi(\Omega) = I$ is locally isochoric at the boundary in the sense of (58), and that the grid is orthogonal at the boundaries except at the corners. Then (64) is strictly stable.*

Proof:
Since the transformation is locally isochoric at the boundary we get $Pv = v$. Thus, the energy method implies

$$\frac{d}{dt}\langle v, v \rangle_h = 2 \sum_{i,k=1}^{2} (v, D_k J^{-1} M_{ik} \tilde{D}_i v)_h$$

Summation by parts yields (v is assumed to have compact support)

$$\frac{d}{dt}\langle v, v \rangle_h = -2 \sum_{i=1}^{2} h_2 \sum_{l=0}^{\nu} \sigma_l v_{0l} (J^{-1} M_{i1} \tilde{D}_i v)_{0l} - 2 \sum_{i=1}^{2} h_1 \sum_{l=0}^{\nu} \sigma_l v_{l0} (J^{-1} M_{i2} \tilde{D}_i v)_{l0}$$
$$-2 \sum_{i,k=1}^{2} (M_{ik} D_k v, J^{-1} \tilde{D}_i v)_h$$

Obviously,

$$\sum_{i,k=1}^{2} (M_{ik} D_k v, J^{-1} \tilde{D}_i v)_h = \sum_{i=1}^{2} \langle \tilde{D}_i v, \tilde{D}_i v \rangle_h$$

Next we turn our attention to the boundary terms. We have

$$\sum_{i=1}^{2} (J^{-1} M_{i1} \tilde{D}_i v)_{0l} = J_{0l}^{-1} |\partial \xi_1|_{0l} \left(|\partial \xi_1|_{0l} (D_1 v)_{0l} + \frac{(\partial \xi_1 \cdot \partial \xi_2)_{0l}}{|\partial \xi_1|_{0l}} (D_2 v)_{0l} \right)$$

The parenthetical expression is recognized as a discretization of the normal derivative (cf. (54)). The other boundary is treated analogously. At $\xi_1 = 0$ we thus define a "normal difference" operator $\tilde{D}_{n_1}$ through

$$(\tilde{D}_{n_1} v)_{0l} = - \left(|\partial \xi_1|_{0l} (D_1 v)_{0l} + \frac{(\partial \xi_1 \cdot \partial \xi_2)_{0l}}{|\partial \xi_1|_{0l}} (D_2 v)_{0l} \right)$$

with a similar definition of $\tilde{D}_{n_2}$ at $\xi_2 = 0$. From (52) it follows that $J_{0l}^{-1} |\partial \xi_1| = |x_{\xi_2}|$. Hence,

$$\frac{d}{dt}\langle v, v \rangle_h = 2 \langle v, \tilde{D}_n v \rangle_\Gamma - 2 \sum_{i=1}^{2} \langle \tilde{D}_i v, \tilde{D}_i v \rangle_h$$

Using $v = Pv$ and the orthogonality assumptions it follows that

$$(\tilde{D}_{n_1} v)_{0l} = -|\partial \xi_1|_{0l} L_{1l}^T v = 0 \quad (\tilde{D}_{n_2} v)_{l0} = -|\partial \xi_2|_{l0} L_{2l}^T v = 0 \quad l > 0$$

At the origin we have

$$(\tilde{D}_{n_1} v)_{00} = L_{0n_1}^T v = 0 \quad (\tilde{D}_{n_2} v)_{00} = L_{0n_2}^T v = 0$$

43

where $L_{0n_1}^T v$ vanishes because of the construction of P; $L_{0n_2}^T v$ disappears since we have shown that L_{0n_2} belongs to the column space of P (cf. proposition 4.2). Hence the boundary sum is identically zero, which proves the proposition. □

Remark: It would still be possible to prove strict stability, even if the grid were not orthogonal at the boundary. To compensate for the loss of orthogonality it is necessary to require that the grid be *globally* isochoric in a neighborhood of the boundary Γ.

5.3   General Parabolic Systems

When considering parabolic equations in general, tangential derivatives may appear in the boundary integrals, potentially calling for integration by parts once more. The occurrence of tangential derivatives depends on the coefficients of the original equation, the geometry, and the presence of mixed derivatives. These criteria are not independent of one another. The following simple example will illustrate this interdependence. Consider the parabolic model equation

$$u_t = u_{x_1 x_1} + u_{x_1 x_2} + u_{x_2 x_2} \quad x \in \Omega \tag{65}$$

where Ω is diffeomorphic to the unit square; the boundary conditions are yet to be specified. The energy method gives (the cross term is integrated with respect to x_1)

$$\frac{d}{dt}||u||^2 = 2 \int_\Gamma (uu_n + n_1 uu_{x_2})ds - 2 \int_\Omega (u_{x_1}u_{x_1} + u_{x_1}u_{x_2} + u_{x_2}u_{x_2})dx$$

The normal and tangential derivatives are defined as

$$
\begin{aligned}
\frac{\partial}{\partial n} &= n_1 \frac{\partial}{\partial x_1} + n_2 \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} &= n_1 \frac{\partial}{\partial n} + \tau_1 \frac{\partial}{\partial \tau} \\
&\quad\quad\quad\quad\quad \Longleftrightarrow & & \\
\frac{\partial}{\partial \tau} &= \tau_1 \frac{\partial}{\partial x_1} + \tau_2 \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_2} &= n_2 \frac{\partial}{\partial n} + \tau_2 \frac{\partial}{\partial \tau}
\end{aligned}
$$

where n is the outward unit normal as usual; the unit tangential τ is chosen corresponding to a positive orientation of Γ. Thus

$$
\begin{aligned}
\tau_1 &= -n_2 \\
\tau_2 &= n_1
\end{aligned}
\tag{66}
$$

If, on the other hand, the cross term is integrated with respect to x_2, we obtain

$$\frac{d}{dt}||u||^2 = 2 \int_\Gamma (uu_n + n_2 uu_{x_1})ds - 2 \int_\Omega (u_{x_1}u_{x_1} + u_{x_1}u_{x_2} + u_{x_2}u_{x_2})dx$$

We must show that

$$\int_\Gamma n_1 uu_{x_2} ds = \int_\Gamma n_2 uu_{x_1} ds \tag{67}$$

44

in order for the energy method to be well-defined. Using the definitions above gives

$$\int_\Gamma n_1 u u_{x_2} ds = \int_\Gamma (n_1 n_2 u u_n + n_1 \tau_2 u u_\tau) ds$$

$$\int_\Gamma n_2 u u_{x_1} ds = \int_\Gamma (n_1 n_2 u u_n + n_2 \tau_1 u u_\tau) ds$$

Clearly, (67) will follow iff

$$\int_\Gamma n_1 \tau_2 u u_\tau ds = \int_\Gamma n_2 \tau_1 u u_\tau ds$$

From (66) and $n_1^2 + n_2^2 = 1$ it follows immediately that

$$\int_\Gamma n_1 \tau_2 u u_\tau ds = \int_\Gamma n_2 \tau_1 u u_\tau ds - \int_\Gamma u u_\tau ds$$

Note that the second integral of the right hand side would vanish identically if Γ were smooth. To simplify the analysis it will be supposed that u is supported only in a neighborhood of the lower left corner. Hence, it will be sufficient to consider the boundary portions Γ_1 and Γ_2 corresponding to $\xi_1 = 0$ and $\xi_2 = 0$. Parametrizing Γ in the positive direction gives (cf. (54))

$$\int_\Gamma u u_\tau ds = \frac{1}{2} \int_1^0 -(u^2)_{\xi_2} d\xi_2 + \frac{1}{2} \int_0^1 (u^2)_{\xi_1} d\xi_1$$

Letting $\xi_2 \to -\xi_2$ in the first integral of the right member gives ($\partial/\partial\xi_2 \to -\partial/\partial\xi_2$, $d\xi_2 \to d\xi_2$)

$$\int_\Gamma u u_\tau ds = \frac{1}{2} \int_{-1}^0 (u^2)_{\xi_2} d\xi_2 + \frac{1}{2} \int_0^1 (u^2)_{\xi_1} d\xi_1 = 0$$

and (67) follows. The energy method is thus well-defined, and we have

$$\frac{d}{dt}||u||^2 \le 2 \int_\Gamma ((1 + n_1 n_2) u u_n + n_1^2 u u_\tau) ds - \left(||u_{x_1}||^2 + ||u_{x_2}||^2 \right)$$

The quantity n_1^2 is discontinuous at the corners. Define the jump discontinuity

$$[n_1^2](x) = n_{1R}^2(x) - n_{1L}^2(x)$$

where $n_{1R}^2(x)$ and $n_{1L}^2(x)$ are the left and right limits of n_1^2 at x (according to the positive orientation of Γ). Straightforward computations show that

$$\int_\Gamma n_1^2 u u_\tau ds = \frac{1}{2} \sum_{i=1}^4 [n_1^2](x_{c_i}) u^2(x_{ci}, t) - \frac{1}{2} \int_\Gamma (n_1^2)_\tau u^2 ds$$

where x_{ci}, $i = 1, \dots, 4$ are the corner points. Thus,

$$\frac{d}{dt}||u||^2 \le \sum_{i=1}^4 [n_1^2](x_{c_i}) u^2 + \int_\Gamma \left(2(1 + n_1 n_2) u u_n - (n_1^2)_\tau u^2 \right) ds - \left(||u_{x_1}||^2 + ||u_{x_2}||^2 \right)$$

45

From this inequality it is obvious that giving Dirichlet data at the corners and Neumann data at the remaining boundaries would yield an energy estimate. In fact, we could even allow inhomogeneous Dirichlet data at the corners and still obtain an energy estimate in terms of the data. The effect of the corners disappears iff $[n_1^2](x) = 0$, which happens iff

(i) $n_{1L}(x) = n_{1R}(x)$

(ii) $n_{1L}(x) = -n_{1R}(x)$

The first case implies that the normal is continuous, i. e., x is not a corner point. The second case is more interesting, since the normal is discontinuous, but the effect on the energy estimate disappears. This illustrates how the geometry can interact with the cross terms. The simplest example is obtained by solving (65) on Ω being the square with $(1, 0)$, $(0, 1)$, $(-1, 0)$, and $(0, -1)$ as vertices. Evidently, the second case holds at the corners, and no corner values should appear in the energy estimate. This can also be seen by a change of coordinates:

$$\begin{aligned}
\xi_1 &= \frac{1}{\sqrt{2}} x_1 + \frac{1}{\sqrt{2}} x_2 \\
\xi_2 &= -\frac{1}{\sqrt{2}} x_1 + \frac{1}{\sqrt{2}} x_2
\end{aligned}$$

Eq. (65) is then transformed into

$$u_t = \frac{3}{2} u_{\xi_1 \xi_1} + \frac{1}{2} u_{\xi_2 \xi_2} \quad \xi \in (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) \times (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$$

The cross term has vanished; instead the equation has become anisotropic. Working in this coordinate system it is apparent that no tangential derivatives — and consequently no point values — will appear when deriving the energy estimate.

To solve (65) by means of finite difference methods it is rewritten in self-adjoint form:

$$\left(J^{-1} u \right)_t = \sum_{k=1}^{2} \left(J^{-1} \left(1 + n_1^{(k)} n_2^{(k)} \right) \partial \xi_k \cdot \partial u \right)_{\xi_k} + \sum_{k \neq l}^{2} (-1)^l \left(n_1^{(k)} n_1^{(k)} u_{\xi_l} \right)_{\xi_k} \qquad (68)$$

where $n^{(k)} = -\partial \xi_k / |\partial \xi_k|$. This equation is discretized in space the usual way. The cross terms

$$(-1)^l \left(n_1^{(k)} n_1^{(k)} u_{\xi_l} \right)_{\xi_k}$$

must be integrated twice to eliminate the tangential derivatives. In the semi-discrete case this amounts to performing summation by parts twice, the second of which will require the introduction of a commutator, thereby obliterating strict stability (except for the second order accurate difference operator). To restore strict stability it would be tempting to reformulate the critical terms in skew-symmetric form:

$$n_1^{(k)} n_1^{(k)} u_{\xi_l} = \frac{1}{2} \left(n_1^{(k)} n_1^{(k)} u \right)_{\xi_l} + \frac{1}{2} n_1^{(k)} n_1^{(k)} u_{\xi_l} - \frac{1}{2} \left(n_1^{(k)} n_1^{(k)} \right)_{\xi_l} u$$

Doing so, however, would introduce lower order energy terms $\langle, \rangle_h$, whose presence would destroy strict stability. The simplest way to resolve this ambiguity is to assume homogeneous Dirichlet data, in which case the boundary terms vanish identically, and (68) would be the preferred choice. The choice of homogeneous Dirichlet data to eliminate the influence of the mixed derivatives arises naturally when solving the Navier-Stokes equations, since at solid boundaries we have zero velocity, and since the cross terms involve only the velocity components.

We now turn to general parabolic systems subject to homogeneous Dirichlet conditions. For general domains Ω eq. (46) is written as

$$
\begin{aligned}
\left(J^{-1} u \right)_t &= \sum_{i,j,k=1}^{2} \left(J^{-1} \frac{\partial \xi_k}{\partial x_i} A_{ij} u_{x_j} \right)_{\xi_k} - \sum_{j=1}^{2} J^{-1} \mathrm{div}(A_j) u_{x_j} + J^{-1} F \quad x \in \Omega \\
u(x,0) &= f(x) \\
u(x,t) &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad x \in \Gamma
\end{aligned}
\tag{69}
$$

where $\mathrm{div}(A_j) = (A_{1j})_{x_1} + (A_{2j})_{x_2}$. Define $C_j \equiv \mathrm{div}(A_j)$. The semi-discrete system is then given by

$$
(J^{-1} v)_t = P \left(\sum_{i,j,k=1}^{2} D_k J^{-1} M_{ik} A_{ij} \tilde{D}_j v - \sum_{j=1}^{2} J^{-1} C_j \tilde{D}_j v + J^{-1} F \right)
\tag{70}
$$

The projection operator P represents the homogeneous Dirichlet conditions.

Proposition 5.3 *The approximation (70) is strictly stable.*

Proof:
Left to the reader $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

6 Inhomogeneous Boundary Conditions

The principle for handling inhomogeneous boundary data is best illustrated by means of a simple example. Consider the one-dimensional advection equation

$$
\begin{aligned}
u_t + u_x &= 0 \qquad x \in (0,1) \\
u(x,0) &= f(x) \\
u(0,t) &= g(t)
\end{aligned}
\tag{71}
$$

The corresponding semi-discrete system reads

$$
\begin{aligned}
v_t + PDv &= (I - P)\tilde{g}_t \\
v(0) &= f
\end{aligned}
\qquad
P = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}
\qquad
\tilde{g} = \begin{pmatrix} g \\ g_1 \\ \vdots \\ g_\nu \end{pmatrix}
\tag{72}
$$

where g_j, $j = 1, \ldots, \nu$ are to be determined later. Obviously, $v_0(t) = g(t)$ if $f_0 = g(0)$ (cf. proposition 2.4). The boundary condition is thus fulfilled at all time. According to proposition 2.4 one has $v = Pv + (I - P)\tilde{g} \iff (I - P)(v - \tilde{g}) = 0$. Hence, the energy method gives

$$\frac{d}{dt}||v||_h^2 = -2(v - (I - P)\tilde{g}, Dv)_h + 2(v, (I - P)\tilde{g}_t)_h$$

Subtracting $2(\tilde{g}, v_t)_h$ from both sides we get

$$2(v - \tilde{g}, v_t)_h = -2(v - \tilde{g}, Dv)_h - 2(\tilde{g}, v_t + PDv - (I - P)\tilde{g}_t)_h + 2(v - \tilde{g}, (I - P)\tilde{g}_t)_h$$

Using (72) and $(I - P)(v - \tilde{g}) = 0$ shows that

$$2(v - \tilde{g}, v_t)_h = -2(v - \tilde{g}, Dv)_h$$

i. e.,

$$2(v - \tilde{g}, (v - \tilde{g})_t)_h = -2(v - \tilde{g}, D(v - \tilde{g}))_h - 2(v - \tilde{g}, \tilde{g}_t + D\tilde{g})_h$$

If $\tilde{g}$ solves the auxiliary problem

$$\begin{aligned} \tilde{g}_t + D\tilde{g} &= 0 \\ \tilde{g}(0) &= f \end{aligned} \tag{73}$$

then

$$\frac{d}{dt}||v - \tilde{g}||_h^2 = -2(v - \tilde{g}, D(v - \tilde{g}))_h = (v_0 - g)^2 - (v_\nu - g_\nu)^2 \leq 0$$

since $v_0 = g$. Thus,

$$||v(t) - \tilde{g}(t)||_h \leq ||v(0) - \tilde{g}(0)||_h = 0$$

Consequently,

$$v(t) = \tilde{g}(t), \quad t \geq 0$$

If $\tilde{g}$ satisfies (73) we get the energy estimate

$$\frac{d}{dt}||\tilde{g}||_h^2 = -2(\tilde{g}, D\tilde{g})_h = g^2 - g_\nu^2$$

Hence,

$$||\tilde{g}(t)||_h^2 + \int_0^t g_\nu^2(\tau)d\tau = ||f||_h^2 + \int_0^t g^2(\tau)d\tau$$

Finally, $v = \tilde{g}$ implies

$$||v(t)||_h^2 + \int_0^t v_\nu^2(\tau)d\tau = ||f||_h^2 + \int_0^t g^2(\tau)d\tau$$

which is identical to the continuous estimate.

It remains to be verified under what circumstances $\tilde{g}$ solves the auxiliary problem (73). Let $\hat{g} = (\begin{array}{cccc} g_0 & g_1 & \cdots & g_\nu \end{array})^T$ be the solution to (73). Hence

$$D^j \hat{g}(0) = D^j f \quad j = 0, 1, \ldots$$

48

and

$$\frac{\partial^j \hat{g}}{\partial t^j} + (-1)^{j+1} D^j \hat{g} = 0, \quad t \geq 0, \quad j = 0, 1, \ldots$$

i.e., for $t = 0$ we get the compatibility conditions

$$\frac{\partial^j \hat{g}}{\partial t^j}(0) + (-1)^{j+1} D^j f = 0, \quad j = 0, 1, \ldots$$

Thus, if we require that the initial-boundary data f and g satisfy

$$\frac{\partial^j g}{\partial t^j}(0) + (-1)^{j+1} (D^j f)_0 = 0, \quad j = 0, 1, \ldots$$

it follows that

$$\frac{\partial^j g}{\partial t^j}(0) = \frac{\partial^j g_0}{\partial t^j}(0) \quad j = 0, 1, \ldots$$

By virtue of being the solution to (73) $g_0(t)$ is analytic in t. Hence, taking the boundary data $g(t)$ to be analytic these equalities imply that $g(t) = g_0(t)$, $t \geq 0$, i.e., $\tilde{g} = \hat{g}$, which proves that $\tilde{g}$ indeed solves (73).

In what follows we shall analyze the general case. Consider the ODE-system

$$\begin{aligned} (J^{-1} v)_t &= P R(t, v) + (I - P)(J^{-1} \tilde{g})_t \\ v(0) &= f \end{aligned} \tag{74}$$

with J^{-1} being the inverse Jacobian, and where

$$R(t, v) = G(t, v) + J^{-1} F(t), \quad G(t, 0) = 0$$

This form arises naturally when discretizing a non-linear PDE in space; $\tilde{g}$ represents the boundary data, and F is the forcing function; $G(t, v)$ is the discretization of the differential operator. It should be pointed out that most operators G occurring in practice are *autonomous*, i. e., $G = G(v)$. We use the tilde notation to emphasize that $\tilde{g}$ is only partially determined, that is, some components are determined by the boundary data of the underlying PDE; the remaining components are unknown. It is no restriction to assume that $\tilde{g} = (\, g_0 \quad \ldots \quad g_\nu \,)^T$ with g_i, $i = 0, \ldots, s$ being the known components. Otherwise, $\tilde{g}$ could be brought to this form by permuting the dependent variables appropriately. As usual we require that P and J^{-1} commute, which is true if the grid is locally isochoric at the boundary. Next, we define the auxiliary problem

$$\begin{aligned} (J^{-1} w)_t &= R(t, w) \\ w(0) &= f \end{aligned} \tag{75}$$

Any solution to (75) will satisfy

$$\frac{d^j}{dt^j}\left(J^{-1} w\right) = R_j(t, w), \quad j = 0, 1, \ldots$$

49

where R_j is defined recursively by

$$R_j(t,w) = \frac{\partial R_{j-1}}{\partial t}(t,w) + \frac{\partial R_{j-1}}{\partial w}(t,w)R(t,w), \quad j = 1,2,\ldots$$
$$R_0(t,w) = w$$

Consequently, at $t = 0$ we have

$$\frac{d^j}{dt^j}\left(J^{-1}w\right)(0) = R_j(0,f), \quad j = 0,1,\ldots$$

Assumption 6.1 *The boundary data $g_i(t)$, $i = 0,\ldots,s$, the initial data f, and the forcing function F satisfy the compatibility conditions*

$$\frac{d^j}{dt^j}\left(J^{-1}\tilde{g}\right)_i(0) = (R_j(0,f))_i, \quad i = 0,\ldots,s, \quad j = 0,1,\ldots$$

If R is sufficiently well-behaved, in particular if G is linear and autonomous, then $w(t)$ will be analytic for $0 \le t < T$. Thus, if we require that the boundary data $g_i(t)$, $i = 0,\ldots,s$, be analytic it follows that

$$g_i(t) = w_i(t), \quad i = 0,\ldots s$$

Furthermore, the unknown components g_i, $i > s$ are of course taken to be

$$g_i(t) = w_i(t), \quad i > s$$

Hence, $\tilde{g} = w$ solves (75).

Remark: It suffices to consider g_i piecewise analytic, since the process can be repeated at $t = t_1$, where t_1 is the critical time when analyticity is lost.

Proposition 6.1 *Let v be the solution to (74) and suppose that assumption 6.1 holds. If the boundary data g_i, $i = 0,\ldots,s$ are piecewise analytic, then*

$$\langle v - \tilde{g}, (v - \tilde{g})_t \rangle_h = (v - \tilde{g}, G(t,v) - G(t,\tilde{g}))_h$$

Proof:
Using (74) and $Pv = v - (I - P)\tilde{g}$, which is true since P and J^{-1} commute, it follows readily that

$$\langle v, v_t \rangle_h = (v - \tilde{g}, R(t,v))_h + (\tilde{g}, PR(t,v))_h + (v, (I - P)(J^{-1}\tilde{g})_t)_h$$

Hence,

$$\langle v - \tilde{g}, v_t \rangle_h = (v - \tilde{g}, R(t,v))_h + (\tilde{g}, -(J^{-1}v)_t + PR(t,v))_h + (v, (I - P)(J^{-1}\tilde{g})_t)_h$$

50

or, using (74),

$$\langle v - \tilde{g}, v_t \rangle_h = (v - \tilde{g}, R(t, v))_h + (v - \tilde{g}, (I - P)(J^{-1}\tilde{g})_t)_h$$

But

$$(v - \tilde{g}, (I - P)(J^{-1}\tilde{g})_t)_h = ((I - P)(v - \tilde{g}), (J^{-1}\tilde{g})_t)_h = 0$$

and so

$$\langle v - \tilde{g}, v_t \rangle_h = (v - \tilde{g}, R(t, v))_h$$

which in turn is equivalent to

$$\langle v - \tilde{g}, (v - \tilde{g})_t \rangle_h = (v - \tilde{g}, R(t, v) - R(t, \tilde{g}))_h - (v - \tilde{g}, (J^{-1}\tilde{g})_t - R(t, \tilde{g}))_h$$

The assumptions on $\tilde{g}$ imply that $(J^{-1}\tilde{g})_t = R(t, \tilde{g})$, which proves the proposition. $\square$

6.1 One-Dimensional Parabolic Systems

We consider the parabolic system (27) with the lower order terms omitted, i. e.,

$$
\begin{aligned}
u_t &= (Au_x)_x + F \\
u(x, 0) &= f \qquad\qquad g(t) = \begin{pmatrix} g^I(t) \\ 0 \end{pmatrix} \\
L_0 u(0, t) &+ L_1 u_x(0, t) = g(t)
\end{aligned}
\tag{76}
$$

The omission of lower order terms is done for convenience only. The boundary data $g(t)$ is assumed to be piecewise analytic in t. The corresponding semi-discrete system reads

$$
\begin{aligned}
v_t &= P\left(DADv + F\right) + (I - P)\tilde{g}_t \\
v(0) &= f
\end{aligned}
\tag{77}
$$

where $\tilde{g}$ satisfies $L^T \tilde{g} = g$, with L given by (18). According to proposition 2.4 we have $(I - P)(v - \tilde{g})$ or, equivalently, $L^T v = L^T \tilde{g}$. Thus

$$L_0 v_0 + L_1 (Dv)_0 = g$$

which shows that the analytic boundary conditions are satisfied to some order of accuracy.

Proposition 6.2 *Suppose that assumption 6.1 holds. Then (77) is strictly stable.*

Proof:
We know that $\tilde{g}$ solves the auxiliary problem

$$
\begin{aligned}
\tilde{g}_t &= G(t, \tilde{g}) + F(t) \\
\tilde{g}(0) &= f
\end{aligned}
$$

where $G(t, \tilde{g}) = DAD\tilde{g}$. Proposition 6.1 then yields (using $J = I$)

$$(v - \tilde{g}, (v - \tilde{g})_t)_h = (v - \tilde{g}, DAD(v - \tilde{g}))_h \leq -(v_0 - g_0)^T A(D(v - \tilde{g}))_0 - 2\delta||D(v - \tilde{g})||_h^2$$

Since $L^T(v - \tilde{g}) = 0$ it follows that

$$L_0^{II}(v_0 - g_0) = 0$$

Furthermore, decompose $v_j - g_j = (v_j' - g_j') + (v_j'' - g_j'')$, where $v_j', g_j' \in \ker(L_1^I)$, $v_j'', g_j'' \in \ker(L_1^I)^\perp$. According to assumption 3.1 we then obtain

$$(v - \tilde{g}, (v - \tilde{g})_t)_h \leq -(v_0 - g_0)^T A(D(v'' - \tilde{g}''))_0 - 2\delta||D(v - \tilde{g})||_h^2$$

Arguing exactly as in the proof of proposition 3.2 gives

$$(D(v'' - \tilde{g}''))_0 = -\tilde{L}_1^{-1} L_0(v_0 - g_0)$$

i. e.,

$$(v - \tilde{g}, (v - \tilde{g})_t)_h \leq \gamma|v_0 - g_0|^2 - 2\delta||D(v - \tilde{g})||_h^2$$

By means of the Sobolev inequality 2.6 we thus arrive at

$$(v - \tilde{g}, (v - \tilde{g})_t)_h \leq (\alpha + \mathcal{O}(h))||v - \tilde{g}||_h^2$$

Hence

$$||v(t) - \tilde{g}(t)||_h \leq e^{(\alpha + \mathcal{O}(h))t}||v(0) - \tilde{g}(0)||_h = 0$$

which is equivalent to $v(t) = \tilde{g}(t)$, $t \geq 0$. To get the final estimate we consider the auxiliary problem. One obtains

$$\frac{d}{dt}||\tilde{g}||_h^2 \leq -2g_0^T A(D\tilde{g}'')_0 - 4\delta||D\tilde{g}||_h^2 + ||\tilde{g}||_h^2 + ||F||_h^2$$

Now

$$L_0 g_0 + L_1(D\tilde{g}'')_0 = g$$

and so

$$(D\tilde{g}'')_0 = -\tilde{L}_1^{-1} L_0 g_0 + \tilde{L}_1^{-1} g$$

Thus

$$-2g_0^T A(D\tilde{g}'')_0 = 2g_0^T A\tilde{L}_1^{-1} L_0 g_0 - 2g_0^T A\tilde{L}_1^{-1} g \leq \gamma|g_0|^2 + |g|^2$$

where the algebraic inequality $2xy \leq \epsilon x^2 + \epsilon^{-1} y^2$ was used. This leads to

$$\frac{d}{dt}||\tilde{g}||_h^2 + |g_0|^2 \leq (\gamma + 1)|g_0|^2 - 4\delta||D\tilde{g}||_h^2 + ||\tilde{g}||_h^2 + |g|^2 + ||F||_h^2$$

The coefficients of this estimate are exactly the same as those of the corresponding analytic inequality. Using $\tilde{g} = v$ and eliminating the boundary terms of the right member by means of the Sobolev inequality yields

$$\frac{d}{dt}||v||_h^2 + |v_0|^2 \leq (\alpha + \mathcal{O}(h))||v||_h^2 + |g|^2 + ||F||_h^2$$

where α is the same constant as in the analytic estimate. Finally, integration with respect to time results in

$$||v||_h^2 + \int_0^t |v_0(\tau)|^2 d\tau \leq e^{(\alpha+\mathcal{O}(h))t} \left(||f||_h^2 + \int_0^t \left(|g(\tau)|^2 + ||F(\tau)||_h^2 \right) d\tau \right)$$

which is the desired estimate. □

Remark: The boundary conditions are used twice — first in conjunction with proposition 6.1 to show that $\tilde{g} = v$, and second with the auxiliary problem to get the actual estimate.

6.2   Two-Dimensional Symmetric Hyperbolic Systems

We consider (the lower order terms are omitted for convenience)

$$\left(J^{-1}u \right)_t = \frac{1}{2} \sum_{i=1}^{2} \left(\left(J^{-1}B_i u \right)_{\xi_i} + J^{-1}B_i u_{\xi_i} \right) + J^{-1}F \tag{78}$$

where

$$B_i = \partial \xi_i \cdot A = \frac{\partial \xi_i}{\partial x_1} A_1 + \frac{\partial \xi_i}{\partial x_2} A_2$$

The boundary conditions are given by

$$\varphi_I(x,t) = S(x)\varphi_{II}(x,t) + g(x,t)$$

where φ_I designates the characteristic variables corresponding to the locally ingoing characteristics; $S(x)$ is assumed to be sufficiently small. At the corner $x(0,0)$ we require that

$$\varphi_I(x(0,0),t) = S(x(0,0))\varphi_{II}(x(0,0),t) + g(x(0,0),t), \quad \varphi(x(0,0),t) = Q_i^T u(x(0,0),t)$$

be satisfied for $i = 1,2$, where $Q_i^T(n_1^{(i)}A_1(x(0,0),t) + n_2^{(i)}A_2(x(0,0),t)Q_i$ are diagonal; $n^{(i)}$, $i = 1,2$ are the two normals associated with the corner $x(0,0)$.

At each discrete boundary point $x_{ij} = x(ih_1, jh_2) \in \Gamma$ the boundary conditions are formulated as

$$L(x_{ij})v_{ij} = g(x_{ij},t), \quad x_{ij} \in \Gamma$$

where $L(x_{ij})$ is given by (32). We note that there are two operators $L_i(x_{00})$ corresponding to the two different normals $n^{(i)}$ at x_{00}. Since $L(x_{ij})$ has full rank it follows that there exists a $\tilde{g}_{ij}$ such that

$$L(x_{ij})\tilde{g}_{ij} = g(x_{ij},t), \quad x_{ij} \in \Gamma$$

Hence, the boundary conditions are

$$L(x_{ij})(v_{ij} - \tilde{g}_{ij}) = 0, \quad x_{ij} \in \Gamma$$

or, in global form,

$$L^T(v - \tilde{g}) = 0$$

where $\tilde{g}$ is partially determined by the boundary data. Eq. (78) is discretized as

$$\left(J^{-1}v\right)_t = P\left(\frac{1}{2}\sum_{i=1}^{2}\left(D_i J^{-1} B_i v + J^{-1} B_i D_i v\right) + J^{-1} F\right) + (I - P)\left(J^{-1}\tilde{g}\right)_t \qquad (79)$$
$$v(0) = f$$

where P is the orthogonal projection corresponding to the global operator L^T.

Proposition 6.3 *Suppose that assumption 6.1 holds. Then (79) is strictly stable.*

Proof:
By assumption 6.1 $\tilde{g}$ solves the auxiliary problem

$$\left(J^{-1}\tilde{g}\right)_t = G(t, \tilde{g}) + J^{-1} F(t)$$

where

$$G(t, \tilde{g}) = \frac{1}{2}\sum_{i=1}^{2}\left(D_i J^{-1} B_i \tilde{g} + J^{-1} B_i D_i \tilde{g}\right)$$

Hence, according to proposition 6.1 we have

$$\langle v - \tilde{g}, (v - \tilde{g})_t\rangle_h = \frac{1}{2}\sum_{i=1}^{2}(v - \tilde{g}, \left(D_i J^{-1} B_i + J^{-1} B_i D_i\right)(v - \tilde{g}))_h$$

Summation by parts yields

$$\langle v - \tilde{g}, (v - \tilde{g})_t\rangle_h = \frac{1}{2}\langle v - \tilde{g}, (n_1 A_1 + n_2 A_2)(v - \tilde{g})\rangle_\Gamma$$

But $L^T(v - \tilde{g}) = 0$, whence $v - \tilde{g}$ satisfies the homogeneous boundary conditions. Consequently, (cf. the proofs of propositions 3.1, 4.1, 5.1)

$$\langle v - \tilde{g}, (n_1 A_1 + n_2 A_2)(v - \tilde{g})\rangle_\Gamma \leq -\gamma|v - \tilde{g}|_\Gamma^2 \leq 0$$

Thus, $\tilde{g}(t) = v(t)$, $t \geq 0$.

In the second part of the proof we apply the energy method to the auxiliary problem. Straightforward computations show that

$$\frac{d}{dt}\langle \tilde{g}, \tilde{g}\rangle_h = \langle \tilde{g}, (n_1 A_1 + n_2 A_2)\tilde{g}\rangle_\Gamma + 2\langle \tilde{g}, F\rangle_h$$

Take an arbitrary point x_{0j} on the boundary portion where $\xi_1 = 0$. We must analyze the quadratic form

$$\tilde{g}_{0j}^T \left(n_1^{(1)} A_1 + n_2^{(1)} A_2\right)_{0j} \tilde{g}_{0j}$$

54

We know that $L^T \tilde{g} = g$, where g is the vector representing the analytic boundary data. Define $\varphi_{ij} = Q^T(x_{ij})\tilde{g}_{ij}$. Hence,

$$(\varphi_{ij})_I = S(x_{ij})(\varphi_{ij})_{II} + g_{ij} \tag{80}$$

and the quadratic form is transformed into

$$\tilde{g}_{0j}^T \left(n_1^{(1)}A_1 + n_2^{(1)}A_2\right)_{0j} \tilde{g}_{0j} = \varphi_{0j}^T \Lambda_{0j} \varphi_{0j}$$

Using (80) gives (omitting the spatial subscripts for simplicity)

$$\varphi^T \Lambda \varphi = \varphi_{II}^T \left(\Lambda_{II} + S^T \Lambda_I S\right) \varphi_{II} + 2\varphi_{II}^T S^T \Lambda_I g + g^T \Lambda_I g$$

It is assumed that $\Lambda_{II} \leq -\gamma$ at x_{0j}. For sufficiently small S we thus get

$$\varphi^T \Lambda \varphi \leq -\frac{\gamma}{2}|\varphi_{II}|^2 + (1 + |\Lambda_I|)|g|^2$$

Now, $|\varphi_I| \leq |S||\varphi_{II}| + |g|$. Hence

$$\varphi^T \Lambda \varphi + \frac{\gamma}{4}|\varphi|^2 \leq -\frac{\gamma}{4}|\varphi_{II}|^2 + \frac{\gamma}{4}|\varphi_I|^2 + (1 + |\Lambda_I|)|g|^2 \leq (3 + |\Lambda_I|)|g|^2$$

for S small enough. It should be underscored that this is exactly the same estimate one gets in the continuous case. At each boundary point x_{0j} we have thus established that

$$\tilde{g}_{0j}^T \left(n_1^{(1)}A_1 + n_2^{(1)}A_2\right)_{0j} \tilde{g}_{0j} + \frac{\gamma_{0j}}{4}|\tilde{g}_{0j}|^2 \leq (3 + |(\Lambda_{0j})_I|)|g_{0j}|^2$$

with a similar expression at points x_{i0} corresponding to $\xi_2 = 0$. Letting $\inf(\gamma_{ij}) \equiv \gamma > 0$ we thus obtain

$$\frac{d}{dt}\langle \tilde{g}, \tilde{g}\rangle_h + \frac{\gamma}{4}\langle \tilde{g}, \tilde{g}\rangle_\Gamma \leq (3 + |\Lambda|_\infty)\langle g, g\rangle_\Gamma + 2\langle \tilde{g}, F\rangle_h$$

Finally, identifying $v = \tilde{g}$, integration gives the energy estimate

$$\langle v(t), v(t)\rangle_h + \int_0^t \langle v(\tau), v(\tau)\rangle_\Gamma d\tau \leq Ke^t \left(\langle f, f\rangle_h + \int_0^t \left(\langle g(\tau), g(\tau)\rangle_\Gamma + \langle F(\tau), F(\tau)\rangle_h\right) d\tau\right)$$

which proves the theorem. $\qquad \square$

Remark: Because of the terms $\varphi_{II}^T S^T \Lambda_I g + g^T \Lambda_I g$ the constant K of the energy estimate will in general satisfy $K > 1$, even if no estimate of the boundary terms $\langle v, v\rangle_\Gamma$ is wanted. For $g = 0$, i. e., homogeneous boundary conditions, the critical terms disappear, and one may take $K = 1$ in case no boundary estimate is needed (cf. the remark following the proof of proposition 3.1).

7 Summary and Conclusions

We have demonstrated that for a given finite-dimensional scalar product $(\cdot,\cdot)_h$ any linear discretized boundary condition can be written as an orthogonal projection operator P that satisfies $(u, Pv)_h = (Pu, v)_h$. It should be noted that the projection is well-defined if the corresponding analytic problem is well-posed. For general boundary conditions one may also have to require that the discretization parameter h be small enough (consistency). The projections P , the summation-by-parts property, and proposition 2.4 constitute the main tools needed to obtain an energy estimate for the semi-discrete case. For a large class of problems it has been established that existence of an energy estimate for the continuous problem implies ditto for the semi-discrete system.

In one space dimension we are no longer required to consider restricted full norms

$$
\Sigma = \begin{pmatrix} 1 & & & \\ & \Sigma^{(1)} & & \\ & & 1 & \\ & & & \ddots \end{pmatrix}
$$

which were used in [4] to prove stability for symmetric hyperbolic systems subject to homogeneous boundary conditions. The main result is the stability proof for mixed hyperbolic-parabolic systems subject to general linear boundary conditions. Assuming certain compatibility conditions the result holds for inhomogeneous boundary data. Reformulating the analytic problem it is possible to obtain strict stability, i. e., we have a time stable semi-discrete approximation that is bounded by the the same exponential growth rate (modulo $\mathcal{O}(h)$) as the analytic problem. For the parabolic part the excess growth rate is induced by the discrete Sobolev inequality. Furthermore, for the hyperbolic part we have used assumption 2.1. In particular, strict stability is obtained for diagonal norms and variable coefficient problems, and for general norms and constant coefficient problems. The stability results hold for finite difference approximations of arbitrary order.

In two space dimensions we are forced to consider diagonal norms in order to get summation by parts in both dimensions. Stability of high-order schemes is obtained for general mixed hyperbolic-parabolic initial-boundary value problems. Again, inhomogeneous boundary conditions are allowed, provided certain compatibility conditions prevail. Using a different norm we obtain strict stability for symmetric hyperbolic systems on non-smooth curvilinear domains, where we allow for general inhomogeneous boundary conditions. As for strict stability of parabolic systems, we are limited to homogeneous Dirichlet data. Mixed derivatives and/or variable geometry may account for "tangential differences" that in general cannot be eliminated without ruining strict stability. An exception is the standard second order method. The requirement that the Dirichlet data be homogeneous is already necessary for the continuous problem. All results obtained for two dimensions generalize to higher dimensions.

The methods presented in this report are similar to finite element methods in that

stability for the semi-discrete system follows more or less directly from the corresponding continuous one. There is, however, one major difference: The FEM technique often results in implicit space discretization, whereas the discretized space operators reported in this paper always are explicit.

There are other ways of imposing boundary conditions so as to ensure time stability (strict stability) when using difference operators satisfying a summation-by-parts property. An elegant technique is proposed in [1]. A so called Simultaneous Approximation Term, SAT for short, is added to the semi-discrete scheme. The SAT will act as a penalty function to enforce an approximation of the discrete boundary conditions. In [1] this approach is used to prove time stability for high-order finite difference approximations of one-dimensional constant coefficient hyperbolic systems. Also, it is not necessary to consider identical difference stencils in the interior. A new and interesting class of such difference operators can be found in [2].

Acknowledgment

8   Appendix

Let P be defined by proposition 2.5 with L given by (42) ($\chi = 0.5$) where $h_1 = h_2 = 1$ for convenience. We shall show that $L_{10}^T P = 0$, which will follow if we can prove that $L_{10}^T \Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T = L_{10}^T$. Straightforward computations show that

$$L^T \Sigma^{-1} L = \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{12}^T & D_{22} & D_{23} \\ D_{13}^T & D_{23}^T & D_{33} \end{pmatrix}$$

where

$$D_{11} = D_{33} = \kappa \begin{pmatrix} \dfrac{1}{\sigma_1} & & \\ & \ddots & \\ & & \dfrac{1}{\sigma_r} \end{pmatrix} \qquad D_{12} = D_{23}^T = \tau^{-1} \begin{pmatrix} \dfrac{d_{01}}{\sigma_1} \\ \vdots \\ \dfrac{d_{0r}}{\sigma_r} \end{pmatrix}$$

and

$$D_{13} = \tau^2 D_{12} D_{23} \qquad D_{22} = \frac{1}{2}\left(\frac{\kappa}{\sigma_0} + \frac{d_{00}^2}{\sigma_0^2} \right) \qquad \kappa = \sum_{k=0}^{r} \frac{d_{0k}^2}{\sigma_k} \qquad \tau = \frac{2\sigma_0}{d_{00}}$$

The inverse is given by

$$(L^T \Sigma^{-1} L)^{-1} = \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{12}^T & T_{22} & T_{23} \\ T_{13}^T & T_{23}^T & T_{33} \end{pmatrix}$$

with

$$T_{11} = T_{33} = D_{11}^{-1} + \frac{\mu}{1 - \mu\sigma} D_{11}^{-1} D_{12} D_{12}^T D_{11}^{-1}$$

$$T_{12} = T_{23}^T = \frac{-(\mu - \sigma\tau^4)}{(1 - \mu\sigma)(1 - \sigma\tau^2)} D_{11}^{-1} D_{12}$$

$$T_{13} = T_{13}^T = \frac{\mu - \tau^2}{(1 - \mu\sigma)(1 - \sigma\tau^2)} D_{11}^{-1} D_{12} D_{23} D_{33}^{-1}$$

$$T_{22} = \frac{(\mu - \sigma\tau^4)(1 + \sigma\tau^2)}{(1 - \mu\sigma)(1 - \sigma\tau^2)}$$

and

$$\sigma = D_{12}^T D_{11}^{-1} D_{12}$$
$$\mu = \sigma\tau^4 + \frac{1}{\nu}\left(1 - \sigma\tau^2\right)^2$$
$$\nu = D_{22} - \sigma$$

Obviously, ν is the Schur complement of D_{22}. Let $\tilde{L}$ be defined by (40). Then

$$L_{10}^T \Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T = L_{10}^T \Sigma^{-1} \tilde{L} \begin{pmatrix} R & S \\ S & R \end{pmatrix} \tilde{L}^T$$

where (using $T_{12}^T = T_{23}$ and $T_{11} = T_{33}$)

$$R = \begin{pmatrix} T_{22}/4 & T_{12}^T/2 \\ T_{12}/2 & T_{11} \end{pmatrix} \quad S = \begin{pmatrix} T_{22}/4 & T_{12}^T/2 \\ T_{12}/2 & T_{13} \end{pmatrix}$$

Furthermore,

$$L_{10}^T \Sigma^{-1} \tilde{L} = \begin{pmatrix} \kappa/\sigma_0 & 0 & d_{00}^2/\sigma_0^2 & 2D_{12}^T \end{pmatrix}$$

Using $D_{22} = (\kappa/\sigma_0 + d_{00}^2/\sigma_0^2)/2$,

$$L_{10}^T \Sigma^{-1} \tilde{L} \begin{pmatrix} R & S \\ S & R \end{pmatrix} =$$

$$\begin{pmatrix} \frac{1}{2}D_{22}T_{22} + D_{12}^T T_{12} & D_{22}T_{12}^T + 2D_{12}^T T_{13} & \frac{1}{2}D_{22}T_{22} + D_{12}^T T_{12} & D_{22}T_{12}^T + 2D_{12}^T T_{11} \end{pmatrix}$$

But

$$\frac{1}{2}D_{22}T_{22} + D_{12}^T T_{12} = \frac{1}{2}(D_{12}^T T_{12} + D_{22}T_{22} + D_{23}T_{23}^T) = \frac{1}{2}$$

and

$$D_{22}T_{12}^T + 2D_{12}^T T_{13} = 2\left(D_{12}^T T_{13} + D_{22}T_{12}^T + D_{12}^T T_{11}\right) - \left(D_{22}T_{12}^T + 2D_{12}^T T_{11}\right)$$

Observing that $T_{12}^T = T_{23}$, $D_{12}^T T_{11} = D_{23}T_{33}$ it follows that the first parenthetical expression vanishes. Thus,

$$L_{10}^T \Sigma^{-1} \tilde{L} \begin{pmatrix} R & S \\ S & R \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & D_{22}T_{12}^T + 2D_{12}^T T_{13} & \frac{1}{2} & -\left(D_{22}T_{12}^T + 2D_{12}^T T_{13}\right) \end{pmatrix}$$

Also,

$$D_{22}T_{12}^T + 2D_{12}^T T_{13} = \rho D_{12}^T D_{11}^{-1} \quad \rho = \frac{2\sigma(\mu - \tau^2) - (\sigma + \nu)(\mu - \sigma\tau^4)}{(1 - \mu\sigma)(1 - \sigma\tau^2)}$$

Substituting $1 - \sigma\tau^2 = d_{00}^2/(\kappa\sigma_0)$ and the expression for μ yields

$$\rho = -\frac{\kappa\sigma_0}{d_{00}^2}$$

and so

$$\rho D_{12}^T D_{11}^{-1} = -\frac{1}{2d_{00}} \begin{pmatrix} d_{01} & \ldots & d_{0r} \end{pmatrix}$$

Consequently,

$$L_{10}^T \Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T = \frac{1}{2d_{00}} \begin{pmatrix} d_{00} & -d_{01} & \ldots & -d_{0r} & d_{00} & d_{01} & \ldots & d_{0r} \end{pmatrix} \tilde{L}^T$$

The j th block column of $\tilde{L}^T$ reads

$$c_j^T = \begin{pmatrix} 0 & \ldots & 0 & \sum_{k=0}^{r} d_{0k}e_k^T & 0 & \ldots & 0 & d_{0j}e_0^T & \ldots & d_{0j}e_r^T \end{pmatrix}$$

where the sum is the j:th block element of c_j, $j = 0, \ldots, r$. Hence,

$$\frac{1}{2d_{00}} \begin{pmatrix} d_{00} & -d_{01} & \ldots & -d_{0r} & d_{00} & d_{01} & \ldots & d_{0r} \end{pmatrix} c_j = \delta_{j0} \sum_{k=0}^{r} d_{0k} e_k^T$$

i. e.,

$$L_{10}^T \Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T = L_{10}^T$$

In a similar fashion one shows that L_{20} also satisfies the above equation.

The simplest example is obtained by discretizing the Neumann conditions using the standard divided difference D_+ in both coordinate directions, i. e.,

$$\begin{aligned}
v_{11} - v_{01} &= 0 \\
v_{11} - v_{10} &= 0 \\
(v_{10} - v_{00})/2 + (v_{01} - v_{00})/2 &= 0
\end{aligned} \tag{81}$$

which leads to

$$\Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T = \frac{1}{18} \begin{pmatrix}
16 & -4 & 0 & \ldots & 0 & -4 & -8 & 0 & \ldots & 0 \\
-2 & 14 & 0 & \ldots & 0 & -4 & -8 & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
-2 & -4 & 0 & \ldots & 0 & 14 & -8 & 0 & \ldots & 0 \\
-2 & -4 & 0 & \ldots & 0 & -4 & 10 & 0 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0
\end{pmatrix} \tag{82}$$

Evidently, any vector $v = Pu$, $P = I - \Sigma^{-1} L (L^T \Sigma^{-1} L)^{-1} L^T$, satisfies (81). Furthermore, by (82), $v_{10} - v_{00} = v_{01} - v_{00} = 0$, which also follows directly from (81).

References

[1] M. Carpenter, D. Gottlieb, and S. Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes. Technical Report 93-21, ICASE, NASA Langley Research Center, Hampton, VA 23681-0001, 1993.

[2] M. Carpenter and J. Otto. High-order cyclo-difference techniques: A new methodology for finite differences. Technical report, ICASE, NASA Langley Research Center, Hampton, VA 23681-0001, 1993.

[3] H.-O. Kreiss and J. Lorenz. *Initial-Boundary Value Problems and the Navier-Stokes Equations*, volume 136 of *Pure and Applied Mathematics*. Academic Press, Inc, San Diego, CA, 1989.

[4] H.-O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*. Academic Press, Inc., 1974.

[5] H.-O. Kreiss and G. Scherer. On the existence of energy estimates for difference approximations for hyperbolic systems. Technical report, Dept. of Scientific Computing, Uppsala University, 1977.

[6] P. Olsson. Stable approximation of symmetric hyperbolic and parabolic equations in several space dimensions. Technical Report 138, Dept. of Scientific Computing, Uppsala Univ., Uppsala, Sweden, Dec. 1991.

[7] G. Scherer. Numerical computations with energy estimates schemes. Technical report, Dept. of Scientific Computing, Uppsala Univ., Uppsala, Sweden, Apr. 1977. In PhD thesis by G. Scherer, 1977.

[8] B. Strand. Summation by parts for finite difference approximations for d/dx. *J. Comput. Phys.* To appear.

RIACS